



(11) Publication number : **0 572 262 A2**

(12)

EUROPEAN PATENT APPLICATION

(21) Application number : **93304151.9**

(51) Int. Cl.⁵ : **H04N 7/133**

(22) Date of filing : **27.05.93**

(30) Priority : **28.05.92 US 891507**

(43) Date of publication of application :
01.12.93 Bulletin 93/48

(84) Designated Contracting States :
DE FR GB NL

(71) Applicant : **C-CUBE MICROSYSTEMS, INC.**
1778 McCarthy Boulevard
Milpitas, California 95035 (US)

(72) Inventor : **Galbi, David E.**
373 River Oak Circle
San Jose, California 95134-1961 (US)
Inventor : **Purcell, Stephen C.**
365 Preston Drive
Mountain View, California 94040 (US)
Inventor : **Chai, Eric Chi-Wang**
1455 Yukon Drive
Sunnyvale, California 94087 (US)

(74) Representative : **Freeman, Jacqueline Carol**
W.P. Thompson & Co. High Holborn House
52-54 High Holborn
London WC1V 6RY (GB)

(54) Decoder for compressed video signals.

(57) A decoder for compressed video signals comprises a central processing unit (CPU), a dynamic random access memory (DRAM) controller, a variable length code (VLC) decoder, a pixel filter and a video output unit. The micro-coded CPU performs dequantization and inverse cosine transform using a pipelined data path, which includes both general purpose and special purpose hardware. In one embodiment, the VLC decoder is implemented as a table-driven state machine where the table contains both control information and decoded values.

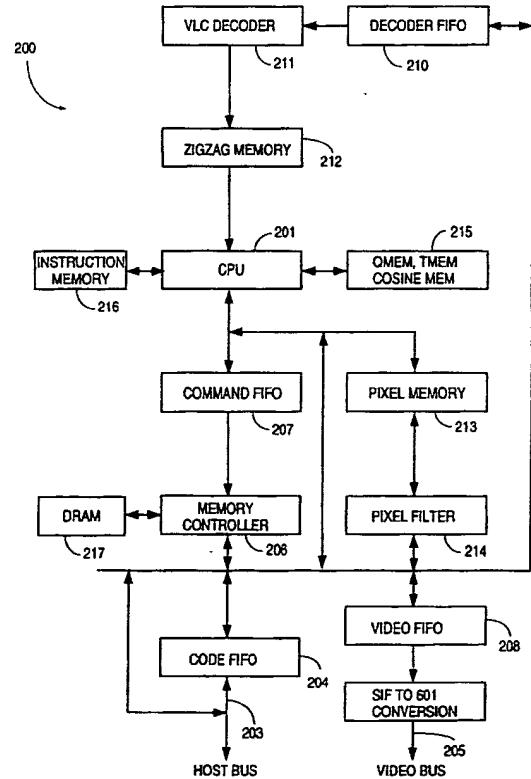


Figure 2

The present invention relates to digital signal processing, and in particular, relates to the decompression of video signals.

Motion pictures are provided at thirty frames per second to create the illusion of continuous motion. Since each picture is made up of thousands of pixels, the amount of storage necessary for storing even a short motion sequence is enormous. As a higher definition image is desired, the number of pixels in each picture is expected to grow also. Fortunately, taking advantage of special properties of the human visual system, lossy compression techniques have been developed to achieve very high data compression without loss of perceived picture quality. (A lossy compression technique involves discarding information not essential to achieve the target picture quality). Nevertheless, the decompression processor is required to reconstruct in real time every pixel of the stored motion sequence.

The Motion Picture Experts Group (MPEG) provides a standard (hereinbelow "MPEG standard") for achieving compatibility between compression and decompression equipment. This standard specifies both the coded digital representation of video signal for the storage media, and the method for decoding. The representation supports normal speed playback, as well as other play modes of color motion pictures, and reproduction of still pictures. The standard covers the common 525- and 625-line television, personal computer and workstation display formats. The MPEG standard is intended for equipment supporting continuous transfer rate to 1.5 Mbits per second, such as compact disks, digital audio tapes, or magnetic hard disks. The MPEG standard is intended to support picture frames of approximately 288 X 352 pixels each at a rate between 24Hz and 30Hz. A publication by the International Standards Organization (ISO) entitled "Coding for Moving Pictures and Associated Audio -- for digital storage media at up to about 1.5Mbit/s," provides in draft form the proposed MPEG standard, which is hereby incorporated by reference in its entirety to provide detailed information about the MPEG standard.

Under the MPEG standard, the picture frame is divided into a series of "Macroblock slices" (MBS), each MBS containing a number of picture areas (called "macroblocks") each covering an area of 16 X 16 pixels. Each of these picture areas is represented by one or more 8 X 8 matrices which elements are the spatial luminance and chrominance values. In one representation (4:2:0) of the macroblock, a luminance value (Y type) is provided for every pixel in the 16 X 16 pixels picture area (in four 8 X 8 "Y" matrices), and chrominance values of the U and V (i.e., blue and red chrominance) types, each covering the same 16 X 16 picture area, are respectively provided in an 8 X 8 "U" matrix and an 8 X 8 "V" matrix. That is, each 8 X 8 U or V matrix covers an area of 16 X 16 pixels. In an-

other representation (4:2:2), a luminance value is provided for every pixel in the 16 X 16 pixels picture area, and two 8 X 8 matrices for each of the U and V types are provided to represent the chrominance values of the 16 X 16 pixels picture area.

5 the 16 X 16 pixels picture area.
The MPEG standard adopts a model of compression
and decompression shown in Figure 1. As shown
in Figure 1, interframe redundancy (represented by
block 101) is first removed from the color motion picture
10 frames. To achieve interframe redundancy removal,
each frame is designated either "intra", "predicted", or "interpolated"
for coding purpose. Intra frames are least frequently provided,
the predicted frames are provided more frequently than the intra
15 frames, and all the remaining frames are interpolated frames.
The compressed video data in an intra frame ("I-picture") is computed only from the pixel values in
the intra frame. In a predicted frame ("P-picture"), only
the incremental changes in pixel values from the last
20 I-picture or P-picture are coded. In an interpolated frame ("B-picture"), the pixel values are coded with
respect to both an earlier frame and a later frame. By
coding frames incrementally, using predicted and interpolated frames,
much interframe redundancy can be eliminated to result in tremendous savings in storage.
Motion of an entire macroblock can be coded by
25 a motion vector, rather than at the pixel level, thereby providing further data compression.

The next steps in compression under the MPEG standard remove intraframe redundancy. In the first step, represented by block 102 of Figure 1, a 2-dimensional discrete cosine transform (DCT) is performed on each of the 8 X 8 values matrices to map the spatial luminance or chrominance values into the frequency domain.

40 Next, represented by block 103 of Figure 1, a process called "quantization" weights each element of the 8 X 8 matrix in accordance with its chrominance or luminance type and its frequency. The quantization weights are intended to reduce to zero many high frequency components to which the human eye is not sensitive. Having created many zero elements in the 8 X 8 matrix, each matrix can now be represented without information loss as an ordered list of a "DC" value, and alternating pairs of a non-zero "AC" value and a length of zero elements following the non-zero value. The list is ordered such that the elements of the matrix are presented as if the matrix is read in a zig-zag manner (i.e., the elements of a matrix A are read in the order A00, A01, A10, A02, A11, A20 etc.). This representation is space efficient because zero elements are not represented individually.

Finally, an entropy encoding scheme, represented by block 104 in Figure 1, is used to further compress the representations of the DC block coefficients and the AC value-run length pairs using variable length codes. Under the entropy encoding scheme, the more frequently occurring symbols are

represented by shorter codes. Further efficiency in storage is thereby achieved.

Decompression under MPEG is shown by blocks 105-108 in Figure 1. In decompression, the processes of entropy encoding, quantization and DCT are reversed, as shown respectively in blocks 105-107. The final step, called "absolute pixel generation" (block 108), provides the actual pixels for reproduction, in accordance to the play mode (forward, reverse, slow motion e.g.), and the physical dimensions and attributes of the display used.

Further, since the MPEG standard is provided only for noninterlaced video signal, in order to display the output motion picture on a conventional NTSC or PAL television set, the decompressor must provide the output video signals in the conventional interlaced fields. Guidelines for decompression for interlaced television signals have been proposed as an extension to the MPEG standard. This extended standard is compatible with the International Radio Consultative Committee (CCIR) recommendation 601 (CCIR-601).

Since the steps involved in compression and decompression, such as illustrated for the MPEG standard discussed above, are very computationally intensive, for such a compression scheme to be practical and widely accepted, the decompression processor must be designed to provide decompression in real time, and allow economical implementation using today's computer or integrated circuit technology.

In accordance with the present invention, an apparatus and a method provide decoding of compressed discrete cosine transform (DCT) coefficients encoded as variable length codes.

In one embodiment, the apparatus comprises a microcoded central processing unit controlling a number of coprocessing units communicating over a global bus. The coprocessing units include (i) a host bus interface unit for receiving a stream of variable length codes, (ii) a memory controller for controlling an external random access memory for storing and retrieving the received stream of variable length codes, (iii) a decompressor and decoder for transforming the compressed variable length codes into DCT coefficients, (iv) an inverse discrete cosine transform processor for transforming the DCT coefficients into pixel values and (v) a pixel filter and motion compensation unit for resampling the pixel values, and for reconstructing the encoded motion sequence based on information in the reference (intra) frames of the motion sequence.

In accordance with another aspect of the present invention, the quantization and the inverse cosine transform functions are performed by special purpose hardware in the central processing unit. In addition, the inverse cosine transform function is performed by a structure comprising (i) a first stage, which receives as operands first, second and third

5 data to compute a result equalling the sum of the first and second data less the third data, and (ii) a second stage, which receives the result from the first stage and compute both the sum and the difference of a fourth datum and the result from the first stage. In one embodiment, the first, second and third data are obtained from a register file, and the results of the first and second stages are returned to the register file, except when the central processing unit directs
10 that the result from the first stage not to be returned ("bypass") to the register file.

In accordance with another aspect of the present invention, the memory controller, which controls a memory system and serves a number of coprocessing units, allocates for each coprocessing unit a first-in-first-out memory so as to separately queue memory access requests for the associated coprocessing unit. A priority circuit in the memory controller grants, under a predetermined priority scheme, memory access to the memory request in the queue having the highest priority. For a memory access request requiring multiple accesses to the memory system, the multiple accesses to the memory system can be preempted by a higher priority memory access request which arrives at the memory controller prior to the completion of the multiple accesses.

In accordance with another aspect of the present invention, the decoding of variable length codes by the decompressor and decoder is controlled by the contents of accessed memory words in a control memory system, such as a read-only memory, which also stores decoded values of the variable length codes. Initially, the decompressor and decoder access the control memory system using an address formed by a predetermined number of bits from the code stream and a predetermined bit pattern according to the command received from the central processing unit. The accessed word in the control memory system is then used to determine if further memory accesses are required. Each word thus accessed contains a decoded value of a variable length code, control information or both. If a further access to the control memory system is necessary, the new access is accomplished using an address formed by a predetermined number of bits obtained from the code stream and a portion of the content of the most recently accessed word in the control memory system. In one embodiment, where a variable length code ("run length") encodes a number of zero values, the accessed words of the control memory system in this decoding method controls the output of these zero values.

55 The present invention is better understood upon consideration of the detailed description below and the accompanying drawings.

Figure 1 is a model of the compression and decompression processes under the MPEG standard.

Figure 2 is a block diagram of a video decompress-

sor circuit 200 in accordance with the present invention.

Figures 3a and 3b show respectively the data flow of CPU 201 and a map of the register file in CPU 201 indicating the contents of registers used in the instruction-cycles of the IDCT operation.

Figure 4 is a flow diagram of one pass of an 8-point IDCT algorithm; in Figure 4, a 1-dimensional IDCT is performed on eight DCT coefficients constituting a row or a column of an 8 X 8 block of DCT coefficients.

Figure 5 shows the sequence of operations of the dequantization and the IDCT operations in CPU 201.

Figure 6a shows circuit 600 which achieves in CPU 201 simultaneous multiplication and butterfly operations in accordance with the present invention.

Figure 6b shows the first stage of circuit 600 in CPU 201's data path.

Figure 7 is a microprogram for computing IDCT in CPU 201, using the IDCT algorithm of Figure 4, in accordance with the present invention.

Figure 8 is a block diagram of a memory controller, showing transfer request fifo (TRF) 207 and DRAM controller 206.

Figure 9 is a block diagram of a pixel filter and motion compensation module comprising pixel filter 213 and pixel memory 214.

Figure 10 is a block diagram of a video interface comprising video FIFO 208 and YUV/RGB conversion circuits.

Figure 11 is a block diagram of a VLC decoder module including VLC decoder 211 and decoder FIFO 208.

A block diagram of an embodiment of the present invention is shown in Figure 2. As shown in Figure 2, a video decoder 200 comprises a central processing unit (CPU) 201, and interfaces with a host computer 202 (not shown) over host bus 203. Host computer 202 provides a stream of compressed video data, which is received by video decoder 200 into a first-in-first-out (FIFO) memory 204 ("code FIFO"). The compressed data received from host computer 202 is decompressed by video decoder 200 and the decompressed video data is provided as video decoder 200's output data over a video bus 205.

Video decoder 200's CPU 201 is a microcoded processor having a control store ("instruction memory") 216. CPU 201 sends commands over a FIFO memory 207 ("command FIFO") to a memory controller 206 ("DRAM controller"), which controls a memory module 217 ("DRAM"). In this embodiment, DRAM 217 comprises dynamic random access memory components, although other suitable memory technologies for implementing a memory system can also be used. DRAM 217 stores both the compressed data received from host computer 202 and the decompressed data for output to video bus 205. The decompressed data for output to video bus 205 are queued

in an output FIFO memory 208 ("Video FIFO").

In this embodiment, the functional modules of video processor 200 communicate over a global bus 209. Control of global bus 209 is granted under a priority scheme to either DRAM controller 206, host bus computer 202, or CPU 201. During operation, compressed video data received from host computer 202 are stored into DRAM 217 by DRAM controller 206 under CPU 201's command. This compressed data is retrieved from DRAM 217 under CPU 201's direction into variable length code (VLC) decoder 211 over a FIFO memory 210 ("decoder FIFO") for decompression. In accordance with the MPEG standard, the decompressed data is reordered by first being stored in "zigzag" order into a memory 212 ("zigzag memory") and then retrieved in row-major order from zigzag memory 212. The row-major ordered decompressed data are then provided to CPU 201 where the decompressed data is "dequantized" and transformed by a 2-dimensional inverse discrete cosine transform (IDCT). The IDCT converts the decompressed DCT coefficients from a frequency domain representation to a spatial domain ("pixel space") representation. In performing the dequantization and IDCT operations, CPU 201 retrieves from a local memory dequantization, cosines and other constants. Temporary storage is also provided by memory unit 215 to store intermediate results of the 2-dimensional IDCT. Memory unit 215 represents a quantization memory 215a, a temporary memory unit 215b, and a cosine memory 215c. The dequantization and the IDCT operations are explained in further detail below.

The pixel space decompressed data are stored into a FIFO memory 213 ("pixel memory"). These pixel space data ("pixels") are filtered and "motion-compensated" by pixel filter 214. The operations of the filtering and motion compensation are discussed in further detail below. Under the direction of CPU 201, DRAM controller 206 stores motion compensated pixels of pixel filter 214 into DRAM 217. The pixels are later retrieved from DRAM 217 by DRAM controller 206, under CPU 201's direction, to provide over global bus 209 to video FIFO 208 as output data of video decoder 200. A CCIR 601 conversion module provides, as a user selectable option, conversion of the decompressed output video data into video data conforming to the CCIR 601 format. The user of the present embodiment can select a 352 X 240 image at a frame rate of 30 frames per second, or a 704 X 240 image at a frame rate of 60 frames per second (i.e. CCIR 601 format). Conversion to the CCIR 601 is achieved by both horizontal interpolation and frame rate conversion techniques.

Internal Global Bus

Global bus 209 is driven from three sources: host computer 202, CPU 201, and DRAM controller 206.

Internal global bus comprises an 8-bit address bus GSEL 209a and a 16-bit data bus GDATA 209b. Two clock periods prior to accessing global bus 209, the unit requesting access asserts a request bit. In accordance with a predetermined priority scheme, in the next clock period, the requesting unit having the highest priority drives onto address bus GSEL 209a the address of the module to which (or from which) the requesting unit desires to send or (receive) data. Separate GSEL addresses are provided for read and write operations. Data are driven onto data bus GDATA 209b by the source of data (i.e. the requesting unit in a write operation, or accessed module in a read operation) in the clock period after the address on address bus GSEL 209a is provided.

In this embodiment, since two clock periods are required per access to DRAM 217, the maximum rate at which data can be written to or requested from DRAM controller 206 is every other clock period.

Host Bus Interface

In this embodiment, code FIFO 204 is 2-bytes wide and holds 32 bytes of compressed code. Host bus interface 203 comprises a 20-bit address bus 203a, a 16-bit data bus 203b, and control signals for performing data transfers or signalling interrupts. Host bus interface 203 includes a host processor clock, a chip clock, a system clock, an address valid strobe (AS) signal, a read/write (R/W) signal, data ready signals UDS and LDS, a reset signal and a test signal. Host computer 202 transfers compressed video data by writing into code FIFO 204 under DMA mode. The compressed video data are transferred on the 16-bit data bus 203b under DMA mode at a rate of 16 bits per bus write cycle. A non-DMA transfer is used by host computer 202 to perform control functions, to access code FIFO 204, and to access DRAM 217 via DRAM controller 206.

Several control registers are accessed by host bus 202 to perform control functions. These registers include: (i) a DMA control register, which allows host computer 202 to enable or disable DMA to code FIFO 204 and to query the status of the code FIFO 204; (ii) a vectored interrupt register, which allows video decoder 200 to perform a vectored interrupt on the host processor; and (iii) a system timer register, which allows host computer 202 to synchronize the compressed video data stream with other MPEG data streams, such as audio.

To access under non-DMA mode any one of these control registers, DRAM 217, or code buffer 204, host computer 202 asserts the "AS," appropriately setting the R/W signal, and places an address on the 20-bit address bus. Bits [20:19] of the 20-bit address and the R/W signal indicate respectively the destination of the access and the access type. For a write access, host computer 202 places

the data to be transferred on the 16-bit data bus and asserts data ready signals UDS and LDS. In response, Video decoder 200 latches the 16-bit data and acknowledges, thereby completing the host bus write cycle. For a read access, video decoder 200 acknowledges the AS signal when the requested data is ready at the 16-bit data bus 203b.

Central Processing Unit

CPU 201 is a microcoded processor having a 24-bit data path and 32 general purpose registers ("register file"). In addition to controlling the co-processors, e.g. memory controller 206, CPU 201 also computes initial addresses for motion compensation (discussed in a later section) and performs dequantization and IDCT. As will be discussed below, both general purpose and special purpose hardware are provided in CPU 201.

The general purpose hardware, which is used by both IDCT and non-IDCT computations, comprises the register file, an arithmetic logic unit (ALU) including a multiplier. The special purpose hardware, which is used in dequantization and IDCT computations, comprises a 5X8 multiplier-subtractor unit 601, a "butterfly" unit 602, cosine read-only memory (ROM) 215c and quantizer memory 215a.

CPU 201 provides special support for the dequantization and IDCT operations specified in the MPEG standard. Specifically, three multiply instructions are designed for these operations. Each multiply instruction also performs simultaneously a "butterfly" computation. A butterfly computation, as it is known in the art, is the simultaneous computation of a sum and a difference of two numbers. Butterfly computations are often encountered in digital filters.

CPU 201 is programmed to perform an IDCT operation in accordance with a 8X8 pixel 2-dimensional IDCT algorithm disclosed in a copending U.S. Patent Application, entitled "A system for Compression and Decompression of Video Data Using Discrete Cosine Transform and Coding Techniques," by Alexandre Balkanski et al., serial No. 07/494,242, filed March 14, 1990, incorporated herein by reference. The IDCT operation is accomplished by performing two 1-dimensional IDCT operations, either row by row or column by column, on an 8 X 8 block or matrix of DCT coefficients. Figure 4 is a flow diagram of the 8-point IDCT algorithm used to operate on one row or one column of the 8 X 8 block. As can be deduced from Figure 4, for each row or column of DCT coefficients, 13 multiplications by a cosine factor and 12 butterfly operations are performed. In Figure 4, the notation C3/16 denotes a multiplication by the cosine factor cosine(3*pi/16), where pi is the well-known mathematical constant 3.14159....

Dequantization of the DCT coefficients are performed in accordance with the MPEG standard. The dequantization coefficients are stored in the quanti-

zation memory ("QMEM") 215a. Dequantization is achieved by multiplying each DCT coefficient in an 8X8 matrix with a corresponding dequantization constant.

The data flow of the dequantization and IDCT operations are summarized in Figure 5. As shown in Figure 5, eight 9-bit DCT coefficients at a time, constituting a row of an 8X8 block of DCT coefficients, are retrieved from zigzag memory 212 at step 501. These DCT coefficients are dequantized at step 502 and multiplied by an appropriate cosine factor at step 503, prior to performing the first 1-dimensional IDCT on the DCT coefficients at step 504. The resulting eight DCT coefficients are then stored in temporary memory 215b at step 505 until the 1-dimensional IDCT is completed on every row of the 8X8 block. Then, eight DCT coefficients at a time, constituting a column of DCT coefficients of the 8 X 8 block, are retrieved at step 506 for the second 1-dimensional IDCT operation. At step 507, the resulting pixel values from the second 1-dimensional IDCT operation are provided to pixel memory 213.

In order to reduce the amount of storage necessary in temporary memory 215b, CPU 201 performs the 1-dimensional IDCT at step 504 alternating between row order for one 8X8 pixel block, and column order for the next 8X8 pixel block. Similarly, the second pass 1-dimensional IDCT operation at step 506 also alternates between column order and row order. Further, for a given 8X8 pixel block, the order in which the 1-dimensional IDCT is performed at step 504 is opposite to the order in which the 1-dimensional IDCT is performed at step 506.

In the present embodiment, the dequantization, the cosine multiply, and the IDCT operations are performed by the same multiplier-subtractor unit of CPU 201. As discussed above, the multiplication instructions of the present embodiment also perform butterfly operations. The present embodiment achieves simultaneous multiplication and butterfly operations using circuit 600 shown in Figure 6a. Circuit 600 of Figure 6a comprises a multiplier-subtractor unit 601 and a butterfly unit 602. As shown in Figure 6a, during a dequantization instruction, dequantization constants are retrieved from quantization memory 215a and each scaled by a 5-bit scaling factor in multiplier 603. The scaled dequantization constant is then provided via multiplexer 604 to multiplier-subtractor unit 601 to be multiplied with the DCT coefficients retrieved from zigzag memory 212. Multiplexer 604 is set to select the dequantization constant during a dequantization instruction and is set to select a cosine factor during a cosine multiply operation. The cosine factor is retrieved from cosine memory 215c. In this embodiment, cosine memory 215c is implemented as a read-only memory.

Prior to arriving at multiplier-subtractor unit 601, in the first stage (shown in Figure 6b) of circuit 600,

each DCT coefficient may be decremented (box 654), made odd, rounded towards zero (box 656) or clipped to a predetermined range (box 658) according to the requirements of the MPEG standard. This first stage of circuit 600, comprising "gate" 651, multiplexer 652, decrementer 653, rounder 656, and clamp 658 are shown in greater detail in Figure 6b.

As shown in Figure 6b, a 9-bit DCT coefficient from zigzag memory 212 can be set to zero by "gate" 651 in response to a control signal "coded". This 9-bit datum, after being zero-padded on the right to form a 14-bit datum, is selected by multiplexer 652 during the execution of a dequantization instruction onto 14-bit bus 681. Alternatively, when executing an instruction other than a dequantization instruction, multiplexer 652 selects the 14-bit datum of bus 682. This 14-bit datum during the execution of an instruction other than a dequantization instruction is the lower order 14 bits of a datum retrieved from a register ("source register") in the register file. The 14-bit datum on bus 681 is decremented by decrementer 653, when required by the MPEG standard, to provide a 14-bit output at the decrementer 653's output terminals. If a decrement operation is not required, the 14-bit datum on bus 681 is provided without modification at decrementer 653's output terminals.

Both bits 0 (LSB) and 4 of the output datum of decrementer 653 can be replaced according to the MPEG standard to provide a 14-bit datum on bus 683. The 14-bit datum on bus 683 can be zeroed by "gate" 656b, if the DCT coefficient from zigzag memory 212 is zero, during execution of a dequantization instruction, or the datum from the register file is zero, during execution of a non-dequantization instruction (e.g. a cosine multiply instruction). Bits 23:19 of the source register is prefixed to the 14-bit output datum of "gate" 656b, and the resulting 19-bit datum is clamped, during execution of a non-dequantization instruction, by clamp 658 to a 14-bit datum having values between -2047 and 2047. Alternatively, during a dequantization instruction, the 14-bit output datum of "gate" 656b is passed through as the output datum of clamp 658. The output datum of clamp 658 is then zero-padded on the right to form a 23-bit datum on bus 684. Multiplexer 659 selects this 23-bit datum to the input terminals of register 660, unless the instruction being executed is a "imac" instruction (see below). During execution of an "imac" instruction, register 660 latches the least significant 23 bits from the source register. The output datum of register 660 is provided as the "X" input datum to multiplier-subtractor unit 601.

Referring back to Figure 6a, multiplier-subtractor unit 601 can, depending on the multiply instruction executed, multiply two numbers X and Y (e.g. in a dequantization, cosine or integer multiply instruction), or compute the value of the expression X*Y-Z (e.g. in an IDCT multiply-subtract instruction). The DCT coefficients are fetched from either zigzag memory 212 or

temporary memory unit 215b to the register file. In addition, the resulting value from an operation in multiplier-subtractor unit 601 can be routed directly as an operand to butterfly unit 602 bypassing the register file.

The butterfly unit 602 computes simultaneously the sum and the difference of its two input operands. Since multiplier-subtractor unit 601 and butterfly unit 602 can each operate on their respective operands during the execution of a multiply instruction, a multiply instruction can result in both a multiplication result and a butterfly result. Additionally, a "pipeline" effect is achieved by using the output value (an "intermediate" result) of multiplier-subtractor unit 601 in a bypass instruction directly in the butterfly operation of the second instruction following the bypass instruction (multiplier-subtractor unit 601 has a latency of two instruction cycles). Under this arrangement, since the intermediate result is not loaded into and then read from a register of the register file, high throughput is achieved.

The results from a butterfly operation of a first pass IDCT are routed into the temporary memory 215b, whereas the results from a butterfly operation of a second pass IDCT operation are "clipped" by clamp 605 and routed to pixel memory 213.

Figure 7 is a microprogram for computing the IDCT provided in Figure 4. In Figure 7, the dequantization, the cosine multiply and the IDCT operations are represented by the instructions shown in Figure 7 as mnemonics "dmac", "cmac" and "imac" respectively. Additionally, instruction "reg(a,b)" assigns the register specified by argument "b" to the name specified in the argument "a". Comments for the microprogram are provided between the "/"* and */" symbols on each line. In the comment portion of the IDCT instructions, the operations performed in the butterfly unit 602 and the multiplier-subtractor unit 601 are set forth respectively under the headings "BUTTERFLYS" and "MULTIPLIES." In Figure 7, in the IDCT portion of the microprogram (i.e. the portion where the instructions imac, dmac and cmac are invoked), the comment lines are numbered from 1-26, indicating the instruction cycles of the loop performing the IDCT.

Figure 7 is read in conjunction with Figures 3a and 3b, which are respectively the data flow through the CPU 201 and a map of contents of registers in the register file. In Figure 3a, each of the rows 1-26 corresponds to the corresponding numbered instruction cycle of the IDCT portion shown in Figure 7. The first two columns, under headings "zmem" and "tmem," shows the operands fetched from zigzag memory 212 and temporary memory 215b respectively. Under the heading "pmem" is shown the result values of the IDCT written to pixel memory 213. The operands under headings "A", "B," and "C" correspond respectively to the operands fetched from the register file to be

provided at the X, Y inputs of butterfly unit 602, and the Z input of multiplier-subtractor unit 601. The value under heading "E" corresponds to the result obtained from the output of the butterfly unit 602. Multiplier-subtractor unit 601 is a 3-stage pipelined multiplier-subtractor. Thus, the values under the headings "E1", "E2" and "E3" correspond to one of the operands in the operation performed respectively by the first, second, third stages of multiplier-subtractor unit 601. Under the heading "R1", "R2" and "R3" are the results from the butterfly unit 602 and multiplier-subtractor unit 602 to be written to the register file. R1 and R2 correspond to the sum and difference results from butterfly unit 602 and R3 corresponds to the result of multiplier-subtractor unit 601.

Figure 3b is a map of the register file, showing the values stored in 16 registers R2-R17 during the instruction cycles of the IDCT portion of the microprogram shown in Figure 7. Each of rows 1-26 shown in Figure 3b shows the contents of registers R2-R17 during the correspondingly numbered instruction cycles of Figure 7.

In Figure 7, the dequantization instruction is represented by the instruction mnemonic "dmac(BTP, r12, r, a, b)", where:

(i) BTP is one of nT, rT, wT, wP, BnT, BrT, BwT and BwP, corresponding respectively to no memory read, read temporary memory 215b, write temporary memory 215b, write pixel memory 213, no memory read with bypass of the register file, read temporary memory 215b with bypass of the register file, write temporary memory 215b with bypass of the register file, and write pixel memory 213 with bypass of the register file;

(ii) "r12" is the address of one of the two registers in which the results of the butterfly computations are stored. Specifically, the register having address r12 stores the sum portion of the butterfly computation, and the register having address r12+1 stores the difference portion of the butterfly computation.

(iii) "r" is the address of the destination register of the dequantization operation, which multiplies the dequantization constant from QMEM 215a to the next DCT coefficient at the output FIFO of zigzag memory 212; and

(iv) "a" and "b" are respectively the source registers of the associated butterfly computation.

In Figure 7, the cosine multiply instruction "cmac" is represented by the instruction mnemonic "cmac(BTP, r12, r, a, b, c)". The arguments to the "cmac" instruction are substantially the same as those described above with respect to the "dmac" instruction. In executing a cosine multiply instruction, a cosine factor, determined by the position of the DCT coefficient in the 8X8 block, is multiplied with the content of the specified register "c" of the register file.

The IDCT instruction is represented in Figure 7 by

the instruction mnemonic "imac(BTP, r12, r, a, b, c)." The arguments to the "imac" instruction are substantially the same as those described above with respect to the "cmac" instruction, except that in executing the imac instruction, a cosine factor is multiplied to the content of the register specified by the argument "c." Before output, the resulting product is subtracted the content of the register specified by the argument "b" in the following instruction.

In Figure 7, the variable names in the IDCT microcode follows the convention now being described. The variable names correspond to those of the values at the nodes of Figure 4, except for names in the form X_i , dX_i , and CX_i , where i is a number from 0 to 7 inclusive. Specifically, X_i refers to a quantized DCT coefficient received from zigzag memory 212, dX_i refers to the value of the DCT coefficient X_i after being dequantized, and CX_i refers to the value of the DCT coefficient X_i after being both dequantized and multiplied by a cosine factor. (Multiplication by a cosine factor is shown as the first step of the IDCT algorithm of Figure 4).

In Figure 7, a name having a suffix "p", e.g. "Ap" on line 3 of the IDCT portion of the microcode shown in Figure 7, denotes a value in the second pass of the 2-dimensional IDCT algorithm. By contrast, a name not having a "p" suffix denotes a value in the first pass of the 2-dimensional IDCT. The results of the first and second passes of the IDCT are passed to temporary memory 215b and pixel memory 213 respectively.

The variable names assigned to the sum and difference results of a butterfly operation are respectively appended the designations "0" and "1." For example, on line 1 of the IDCT portion of the algorithm shown in Figure 7, the comment's expression " $Bp=b(X3p, X5p)$ " explains that the butterfly portion of the "dmac" operation takes the operands $X3p$ and $X5p$ and computes the sum and differences of these operands as, respectively, $B0p$ and $B1p$.

Values that are used as both the input datum at the subtract input of multiplier-subtractor 601 and as an input to the butterfly unit 602 are indicated by the "%" sign in the comment portion of the IDCT operations in Figure 7's microprogram. For example, on line 11 of the IDCT portion, the expression " $iB1p=imac(B1p, %B0p)$ " explains that the operand $B0p$ in the multiplier-subtractor portion of the imac instruction is used in the next instruction. Thus, on line 12, it is shown the expression " $AAp=(A0p, %B0p)$ " to indicate that the $B0p$ is used as an operand in the butterfly portion of the imac instruction on line 12.

Finally, in Figure 7, results of multiplier-subtractor unit 601 which are passed directly to the butterfly unit 602, bypassing the register file, are indicated by the "\$" symbol. For example, on line 4, the expression " $$cX4=cmac(dX4)$ " indicates that the result $cX4$ of the cosine multiply performed on operand $dX4$ is passed directly to butterfly unit 602, bypassing the

register file.

Instruction Memory

5

Instruction memory 216, which stores the microcodes used for executing CPU instructions, comprises a static random access memory (SRAM). Instruction memory 216 is loaded by host computer 202 upon initialization of CPU 201. To effectuate a microcode change, when necessary, the microcodes in the SRAM can be overlayed by microcodes from the DRAM 217.

15

Memory Controller

DRAM controller 206 interacts directly with DRAM 217, generating the signals required to read and write the external DRAM 217. DRAM controller 206 receives from CPU 201 via command FIFO 207 a starting address and offset information. DRAM controller 206 computes subsequent addresses if multiple transfers are requested. In this manner, since CPU 201 need not generate every address for each memory access, CPU 201 is provided more bandwidth for IDCT operations.

Figure 8 is a block diagram of the memory controller module of the present embodiment. As shown in Figure 8, the memory controller module comprises DRAM controller 206 and command FIFO 207, also known as transfer request FIFO (TRF) 207. A pending DRAM access is initiated by CPU 201 writing into memory buffer 801 of TRF 207 an entry indicating (i) the starting address of the DRAM access in bytes, (ii) whether the requested access is a read or write access, (iii) the number ("length") of memory words to be fetched, and (iv) if appropriate, an offset. In the present embodiment, memory buffer 801 holds 11 DRAM access request entries, allocated in order of priority to the following data source or destination: (i) one entry for luminance video FIFO 208a, (ii) one entry for code FIFO 204, (iii) one entry for decoder FIFO 210, (iv) five entries for pixel memory FIFO 213, (v) one entry for either a host memory request or CPU a memory request, and (vi) one entry for chrominance video FIFO 208b. For the purpose of understanding memory controller 206's operation, the entry or entries allocated to each source or destination can be regarded as a FIFO.

50

A DRAM access request becomes pending after (i) CPU 201 writes a memory access request entry into register 804, which is latched into memory buffer 801 and (ii) a request line corresponding to the data source or destination is asserted. A status register file 803a-803f provides for each request line a register to indicate whether or not a memory access request is pending. Since the present embodiment allocates five entries to pixel memory 213, five bits are provided to indicate the number of pending memory access

requests related to pixel memory 213. Naturally, one bit is provided for each of the other status registers corresponding to the remaining data sources or destinations. By asserting a read request signal, the entries in memory buffer 801 can be read by CPU 201. The entry is read by CPU 201 from register 805.

Upon completion of each DRAM access, the length of the memory words to be fetched is decremented by one, except when the 3 least significant bits of the length field of the TRF entry is zero. A zero value in the length field of a TRF entry indicates that the specified offset is to be subtracted from the length after each memory access. Each entry in a TRF entry also indicates the transfertype and whether a byte write (i.e. only 8 of the 16 bits of the memory word are overwritten) is performed.

In this embodiment, host computer 202 can also request DRAM access in substantially the same manner as CPU 201 using a host request line, rather than CPU 201's request line, although host computer only writes into TRF 207 and does not read from TRF 207.

Priority arbitration among pending memory access requests are accomplished by priority arbitration circuit 802 according to the priority scheme set forth above. If DRAM controller 206 is idle, the highest priority request is sent by TRF 207 to DRAM controller 206 by writing into register 806. If a higher priority request is received by TRF 207 while DRAM 206 is processing a lower priority request, priority arbitration circuit 802 sends DRAM controller 206 a "higher priority request pending" signal. In this embodiment, if the current memory access crosses a page boundary while a higher priority request is pending, DRAM controller 206 returns the uncompleted lower priority request back to TRF 207, and begins processing the higher priority request.

TRF 207 issues a "code fifo emptying" request to transfer to DRAM 217 the content of code FIFO 204 when DRAM controller 206 is idle and no DRAM access request is pending at TRF 207. This code fifo emptying request can be issued so long as there is a valid TRF entry corresponding to a request of code FIFO 204 and code FIFO 204 contains at least one or more words, even though code FIFO 204's memory access request line is not asserted. The code fifo emptying request ensures that the last few words of the code FIFO 204 are transferred to DRAM 217.

DRAM controller 206 receives a DRAM access request from register 806, which is written by TRF 207. The format of a DRAM access request in register 806 is the same as the format of a TRF entry in memory buffer 801. Address generation logic 807 calculates subsequent memory addresses based on the starting address, length and offset information of the DRAM access request received. DRAM controller 206 is controlled by state machine 810, which also detects and handles, when the current DRAM access crosses a page boundary, preemption of the incomplete

DRAM access request by another pending DRAM access request in the manner described above.

- 5 When DRAM controller 206 completes a DRAM access, a "memory request done" signal is sent to TRF 207 to allow TRF 207 to allocate the TRF entry in memory buffer 801 to a new request from the same source or destination as the completed DRAM access. In this embodiment, DRAM controller 206 sends an "almost done" signal at the following times: (i) a few cycles prior to completion of the current DRAM access, (ii) a "kill" signal aborting the current access is received from TRF 207, and (iii) a page crossing is expected during the current DRAM access, and a higher priority DRAM access request is pending. When the "almost done" signal is asserted, CPU 201's access to TRF 207 is disabled, to free bus 812 for communication between TRF 207 and DRAM controller 206.
- 10 DRAM controller 206 provides the necessary interface signals to DRAM 217 and controls DRAM 217's refresh activities. A refresh counter keeps tracks of the number of cycles before a refresh is due. If DRAM controller 206 becomes idle prior to the count in the refresh counter reaching zero, a DRAM refresh is performed. Alternatively, when the count in the refresh counter reaches zero, a DRAM refresh is performed after completion of the current DRAM access, or when a page boundary is crossed.
- 15
- 20
- 25
- 30

Variable Length Code (VLC) Decoder

- 35 Like DRAM controller 206 and Pixel filter 214, VLC decoder 211 serves as a slave processor to CPU 201. The instructions of VLC decoder 211 perform the following functions: (i) receive into decoder FIFO 210 under CPU 201's direction a stream of variable length code retrieved from DRAM 217; (ii) decode a variable length code according to the MPEG standard; (iii) construct 8X8 blocks of pixels for "unzigzagging" and dequantization in CPU 201; and (iv) providing up to 15 bits at a time the bits of the code stream.
- 40

- 45 Figure 11 shows VLC decoder module including VLC decoder 211 and decoder FIFO 210. As shown in Figure 11, decoder FIFO 210 receives from DRAM 217 on global bus 209 a stream of variable length codes. Control information (i.e. commands) is also received from CPU 201 and stored in a decoder command register, which is part of global data decode unit 1106. The decoded values of certain variable length codes are provided to zigzag memory 212 on 9-bit zdata bus 1101. Other output values of VLC decoder 211 are provided on global bus 209. A status register (not shown) provides status information which can be accessed by CPU 201 through global bus 209.
- 50
- 55

Commands to VLC decoder 211 are 6-bit wide. When set, bit 5 (i.e. the most significant bit) resets VLC decoder 211. During normal operation, i.e. when bit 5 is zero, the lower 5 bits (4:0) encode either (i) one

of fifteen "get bit" commands, which output 1-15 bits from the code stream to global bus 209, or (ii) the remaining VLC decoder commands. These remaining decoder commands are "mba" (macroblock address), "mtypei" (intraframe macroblock), "mtypep" (predictive frame macroblock), "mtypeh" (h.264 type macroblock), "mtypeb," (bidirectional macroblock), "mv" (motion vector), "cbp" (coded block pattern), "luma" (luminance block), "chroma" (chrominance block) and "non-intra" (block with no dc component)¹. These remaining VLC decoder commands direct decoding by VLC decoder 211 the variable length code at the "head" of the code stream. Except for the "luma", "chroma" and "non-intra" commands, which decoded values are output on zdata bus 1101, the output decoded values of the VLC decoder 211's commands are stored in a decoder register (formed by registers 1102a-d) and provided to CPU 201 on global bus 209.

In this embodiment, decoder FIFO 210 is a 32 X 16-bit FIFO, addressable by 5-bit write and read pointers, which are kept in fifo address logic unit 1103. Freeze logic unit 1104 suspends operation of VLC decoder 211 when decoder FIFO 210 is empty. VLC decoder 211 is controlled by a control store, shown in Figure 11 as 1024 X 15-bit read-only memory (ROM) 1105, which also holds the decoded values of each variable length code. Decoding of variable length codes in ROM 1105 is performed by a table lookup.

In the embodiment shown in Figure 11, if the command in the decoder command register is a "get bits" command (i.e. bit 4 of the command is zero), ROM address generator 1107 generates an address comprising (i) a preassigned bit pattern (in this case, 6-bit value 101011) and (ii) the least significant four bits of the command. Otherwise, if the command is other than a "get bits" command, the 10-bit address comprises (i) two zero leading bits, (ii) the least significant four bits of the command and (iii) the first four bits at the head of the code stream.

Because the variable length codes decoded by VLC decoder 211 can be as long as 12 bits and, as can be seen from the ROM address generated, at most four bits of the code stream are used per access to ROM 1105, decoding a given variable length code may require multiple clocks and multiple accesses to ROM 1105. Other instructions, such as "luma," also require multiple clocks and multiple accesses to ROM 1105 to complete. The most significant bit (14) of the current word of ROM 1105 ("current ROM word"), when set, indicates either execution of the current command is complete, or if the current command is one of the block commands (i.e. either "luma," "chroma," or "non-intra"), a runlength is identified. When a run length is identified, a number of zeroes (equalling the run length identified) are "unpacked" for output on

zdata bus 205. A block command requires the special handling described below.

- 5 Bits 13 and 12 of the current word encodes the number of bits to advance the head of the code stream. In the present embodiment, advancing the head of the code stream are performed by left and right shifters 1109 and 1110 respectively, under the control of bit stream logic 1108.
- 10 When bit 14 of the current ROM word is zero, indicating incomplete execution of the current command, to provide the next ROM address, six bits (9:4) of the current ROM word are combined with either (i) the next four bits at the head of the code stream, when bit 11 of the current ROM word is set, or (ii) another four bits (3:0) of the current ROM word, when bit 10 of the current ROM word is set. The value of bit 11 of the current ROM word indicates that the next ROM access is for decoding purpose, and thus requiring that the remaining four bits of the next ROM address to be taken from the head of the code stream. Alternatively, if the next ROM access is for control purpose, as indicated by the value of bit 10 of the current ROM word, the remaining four bits of the next ROM address is taken from bits 3:0 of the current ROM word.
- 15
- 20
- 25

- 30 If the current command is a block command, and bit 14 of the current ROM word is set, indicating that the run length portion of an encoded AC value-run-length pair is identified, the next ROM address is formed by a predetermined 4-bit pattern (in this embodiment, 4'b1101), and the identified 6-bit run length. The identified runlength is found in (i) zdata counter 1111, if the end of block (EOB) symbol is identified; (ii) the value obtained by cascading the contents of registers 1102b and 1102c, if the short escape symbol is identified; (iii) the value obtained by cascading the contents of registers 1102a and 1102b, if the long escape symbol is identified; or (iv) bits 10:6 of the current ROM word, otherwise. During processing of either a short or long escape symbol, VLC decoder 211 verifies that the 16-bit level code (i.e. the AC value) is within one of the permissible ranges of value. There are three illegal ranges: (i) the value represented by bit pattern 1000_0000_0000_0000, (ii) the range represented by values between 1000_0000_1000_0001 and 1000_0000_1111_1111; and (iii) the range represented by values between 0000_0000_0000_0000 and 0000_0000_0111_1111.
 - 35
 - 40
 - 45
 - 50
 - 55
- The verification that a level code is within a legal range is accomplished by mapping the 4 bits shifted from the code stream every clock period to the low addresses of the ROM, using specific bits of the ROM address last accessed, in the same manner as discussed above with respect to a decoding operation. If the 16-bit level code is within an illegal range, the

¹ Other than "mtypeh," these commands correspond to data objects defined in the MPEG standard. Mtypeh represents a macroblock defined under the h.264 standard which is used in teleconferencing applications.

contents of the address in the ROM reached will signal the illegal 16-bit level code.

In the present embodiment, the output register of decoder FIFO 210, 16-bit register 1112, 5-bit register 1113, 4-bit register 1102d, 4-bit register 1102c, 4-bit register 1102b and 3-bit register 1102a form a 7-stage pipelined data path. In addition, the output data of registers 1102a-1102b can also be treated as a 15-bit register.

At the beginning of each clock period, the left shifter 1109 provides to register 1113 the 5 bits at the head of the code stream. Four of these five bits may be used to access the next ROM word, which provides the number of bits (up to 4 bits) to advance the head of the code stream in the next clock period. In this embodiment, the head of the code stream is monitored by a bit pointer in bit stream logic unit 1108. One clock period after the bit pointer advances (towards the least significant bit) past the code bit at bit 0 of register 1112, the content of the output register of decoder FIFO 210 is loaded into register 1112, and the next entry in decoder FIFO 210 is loaded into the output register of decoder FIFO 210. Because the most significant 9 bits of the output register of decoder FIFO 210 is available to left shifter 1109, 5 bits at the head of the code stream (which is now in the output register of decoder FIFO 210) can be provided by left shifter 1109, without stalling, to form the next ROM address. Although only four bits are used to form the next ROM address, the fifth bit at the head of the code stream is used immediately in a block command after resolving an AC coefficient to determine the sign of the amplitude value to follow.

In addition to providing right shifting to the 5-bit output of left shifter 1109, right shifter 1110 also sign extends the shifted values of the DC and AC components of the luma and chroma block commands.

As discussed above, control of VLC decoder 211 is accomplished by ROM 1105. For example, after decoding the run length of an AC coefficient-run length pair, each ROM word accessed until the end of execution of the current block command will direct decrementing the zdata counter 1111 and enables a zero value to be output on zdata bus 205.

The right and left shifters 1110 and 1109 provide shifting of bits in the "get bits" commands. Since at most 4 bits are shifted per clock period, multiple clock periods are necessary to get more than 4 bits. In the first clock, for a "get n bits" command, (n modulo 4) bits are right shifted and the remaining number of bits are successively shifted 4 bits at a time into the pipeline formed by registers 1102a-1102d.

When the output value of VLC decoder 211 is taken from the current ROM word, bits 14-10 of global bus 209 is set to zero, and bits 9-0 of the current ROM word is output as bits 9-0 on global bus 209 through multiplexers 1114a and 1114b. If the output value of VLC decoder 211 is taken from the code stream,

multiplexers 1114a and 1114b select the output data of register 1102d and right shifter 1110 respectively. Multiplexers 1114a and 1114b can each be selected to provide inverted output values. Such inverted output values are desirable for providing, during execution of a block command, when necessary, 1's complement for a DC amplitude value, or a 2's complement value for an AC amplitude value. Zdata incrementer 1113 completes the 1's complement or 2's complement computation.

Pixel Filter and Motion Compensation

Pixel filter 214 receives reference frames from memory controller 206 and retrieves from pixel memory 213 the decompressed video data from CPU 201. In accordance with the MPEG standard, the reference frames are combined with the decompressed video data using one or more motion vectors, which relates ("predicts") the video data to the reference frames. The resulting video image is written back to DRAM 217 for later output via video FIFO 217 and video bus 205. Under the MPEG standard, the decompressed video data may represent no prediction (i.e. independent of a reference frame), backward prediction (i.e. dependent upon a reference frame of a later time), forward prediction (i.e. dependent upon a reference frame of an earlier time), or interpolated prediction (i.e. dependent upon both a reference frame of an earlier time and a reference frame of a later time).

In the present embodiment, if the video data are not of the "no prediction" type, blocks of one or more reference frames are fetched from DRAM 217. These blocks are each 9X9 components. Since each page of DRAM 217 stores 8 rows and 32 columns of pixels (256 pair of pixels per page), a fetch of a 9X9 block of components crosses at least one page boundary. (In fact, because two pixels are stored in one word of DRAM 217, the actual fetch involves a 10X9 block of pixels). To minimize page crossings, the present embodiment uses the method of access in which all the pixels of the 9X9 block residing in one memory page are accessed before the remaining pixels of the block residing in another memory page are accessed. This method was discussed with respect to motion compensation in an embodiment disclosed in the copending parent application, serial no. 07/669,818, incorporated by reference above.

Figure 9 is a block diagram of pixel filter 213. Pixel pairs are fetched from DRAM 217 and provided to pixel filter 213 on global bus 209. The motion vector consists of x and y components, which are respectively stored in x and y registers (not shown). The x component of the motion vector indicates whether the first column in the 10X9 block of pixels fetched is part of the 9X9 pixel reference block. The y component of the motion vector indicates how many rows of the

10X9 block fetched are in the first memory page.

Every other cycle a pixel pair arrives at global bus 209, and every cycle pixel filter 213 processes one pixel. Column memory 901, which is a 9X8 bit random access memory, stores the last column of pixels previously accessed. As the pixels of the present column arrives, each arriving pixel is averaged (i.e. filtered in the x direction) by adder 902 with the pixel of the same row stored in column memory 901. The arriving pixel then replaces the corresponding pixel stored in column memory 901. The result of adder 902 is latched into pipeline register 903.

The filtered pixels are then averaged (i.e. filtered in the y direction) by adder 905 with the filtered pixels of the previous row stored in row memory 904. Each incoming pixel from pipeline register 903 replaces the corresponding pixel in row memory 904. The resulting filtered pixel from adder 905 are latched successively into pipeline register 906. The net result of the averaging in the x and y direction is a translation ("resampling") of the 8X8 block by one-half pixel, as required by the MPEG standard. The resampled reference frame is then added by adder 906 to the decompressed video data in pixel memory 213. Pixel memory 213 comprises two halves, each half alternately receives in a double buffer scheme decompressed data from CPU 201 and provides pixels to the pixel filtering in pixel filter 213.

In the present embodiment, x and y registers are provided for both forward and backward motion vectors. In processing interpolated predicted blocks, the forward reference frame (associated with the forward motion vector) is fetched first for forward compensation. The result of the forward compensation is stored in pixel memory 213 for backward compensation using the backward reference frame, which is fetched after the forward compensation.

Video Interface

The filtered and motion compensated video data are provided as output, on video bus 205, of video decoder 200 to the video interface. A block diagram of the video interface is shown in Figure 10. As shown in Figure 10, video data are provided as pixel pairs to video interface (generally indicated in Figure 10 by reference numeral 1000) via global bus 209. CPU 201 also provides control information to video interface 1000 over global bus 209. Such control information includes, for example, conversion factors necessary to convert between YUV represented data (i.e. luminance-chrominance representation) and RGB represented data, and the starting and ending positions of active data in a scan line. The conversion factors are stored in registers 1001. Timing logic 1002 which receives synchronization signals VSYNC and HSYNC (vertical and horizontal synchronization signals) synchronizes the operation of the video interface 1000

with the video data stream received.

The pixels in each pair of incoming pixels are YUV represented and are the same Y, U or V type. These pixels are stored in video FIFO 208, which comprises in fact two fifos, respectively referred as video FIFO 208a and video FIFO 208b. Video FIFO 208a and video FIFO 208b store luminance (Y) and chrominance (U or V) data respectively.

In this embodiment, the YUV represented data can be converted for output, at the user's option, as RGB represented data. Conversion from YUV represented data into RGB represented data is accomplished in block 1003. A synchronizer circuit 1004 receiving externally provided video clock signal VCLK provides the output video data on 24-bit bus 1006 at the desired rate.

Claims

1. An apparatus for decoding compressed video data, the compressed video data being discrete cosine transform (DCT) coefficients of pixels compressed into variable length codes, the apparatus comprising:

a global bus;

a central processing unit, the central processing unit controlling over the global bus a plurality of coprocessing units, the plurality of coprocessing units comprising:

(a) a memory controller for controlling storing and retrieving data from a memory system;

(b) a host interface unit for receiving from a host computer the variable length codes and for causing the variable length codes to be stored in said memory system;

(c) a decoder causing said variable length codes to be retrieved from the memory system for decompressing the variable length codes to recover the DCT coefficients;

(d) inverse discrete cosine transform unit receiving said DCT coefficients for transforming the DCT coefficients to recover said pixels; and

(e) pixel output unit for providing the pixels to an image rendering device.

2. An apparatus as claimed in claim 1, wherein the inverse discrete cosine transform unit is part of the central processing unit.

3. An apparatus as claimed in claim 1 or 2, wherein the central processing unit comprises a micro-coded processor.

4. An apparatus as claimed in claim 1, 2 or 3, wherein the variable length codes further comprises encoded motion vectors decoded by said decod-

- er, and wherein the memory system storing reference frames of pixels, the pixel output unit further comprises:
- a motion compensation unit for combining the pixels recovered by the inverse cosine transform unit with the reference frames of pixels in accordance with the motion vectors to reconstitute pixels of an image; and
 - video output bus for providing the pixels of an image in accordance with a predetermined image representation convention to the image rendering device.
5. A structure for computing alternatively a discrete cosine transform or an inverse discrete cosine transform, comprising:
- a multiplier-subtractor circuit receiving first second and third input data for providing a fourth datum which equals the product of the first and second input data less the third input datum; and
 - 20 a "butterfly" circuit receiving the fourth datum and a fifth datum for providing first and second output data which equal respectively the sum and differences of the fourth and fifth data.
6. A structure as claimed in claim 5, further comprising:
- 25 a plurality of registers;
 - means for providing selected constant values; and
 - 30 a control unit for retrieving from the registers and said means the first, second, third and fifth data, for storing the first and second output data in registers, and, in accordance with an algorithm, for alternatively storing the fourth datum in the registers and bypassing the registers.
7. A structure as claimed in claim 5 or 6, further comprising:
- 35 a decrementer for receiving a sixth input datum and providing as output a seventh datum, said seventh datum being one of (i) the sixth datum and (ii) the sixth datum decremented by a predetermined value; and
 - 40 a clamp circuit receiving the seventh datum for and providing as an output datum the first datum, the first datum being the seventh datum restricted to a predetermined range.
8. A structure as claimed in claim 5, further comprising:
- 45 a multiplexer for alternately selecting the first and second output data of the "butterfly circuit"; and means for coupling the output datum selected by the multiplexer to one of (i) a temporary memory for storing intermediate data of the discrete cosine transform, and (ii) a memory for
 - 50 storing results of the discrete cosine transform.
9. A structure as claimed in claim 8, further comprising a clamp circuit for restricting the output datum selected to a predetermined range prior to coupling the output datum selected to the memory for storing results.
10. A structure as claimed in claim 5 or 6, further comprising means for retrieving from a memory a sixth and a seventh datum, and wherein in accordance with a predetermined algorithm, the butterfly circuit accepts as input data the sixth and seventh data in lieu of the fourth and fifth data.
15. A structure as claimed in any one of claims 5 to 10, wherein the multiplier-subtractor comprises a pipelined structure receiving a clock signal, such that, even though the multiplier-subtractor completes an operation over multiple periods of the clock signals, the multiplier-subtractor receives input data every period of the clock signal and provides output data every period of the clock signal.
20. A structure as claimed in any one of claims 5 to 11, wherein each of the discrete cosine transform and the inverse discrete cosine transform is performed in two passes over a square matrix of input data, wherein the input data are provided to the structure in row order in one pass, and provided in column order in the other pass.
25. A structure for controlling a memory system, said structure receiving request for accessing the memory system from a plurality of information processing units, the structure comprising:
- 30 a first-in-first-out buffer associated with each of the information processing units, the buffer queuing the memory access requests of the associated information processing unit;
 - 35 a priority arbitration unit for determining which of the memory access requests has the highest priority, in accordance with a predetermined hierarchy of the information processing units; and
 - 40 a dispatch circuit for causing an access of the memory system in accordance with the highest priority memory request.
45. A structure as claimed in claim 13, wherein one of the memory requests requires multiple accesses to the memory system, the structure further comprising an interrupt circuit for interrupting the multiple accesses to the memory, when a higher priority memory access request arrives prior to completion of the multiple accesses to the memory system, and for causing the higher priority

memory access request to be performed in lieu
of completing the interrupted multiple accesses.

- | | |
|--|----|
| 15. A method for decoding compressed video data,
the compressed video data being discrete cosine
transform (DCT) coefficients of pixels com-
pressed into variable length codes, the method
comprising providing apparatus according to any
one of claims 1 to 4. | 5 |
| 16. A method for computing a discrete cosine trans-
form and an inverse discrete cosine transform,
comprising providing a structure according to any
one of claims 5 to 12. | 10 |
| 17. A method according to claim 16 further compris-
ing the step of restricting said output datum se-
lected to a predetermined range prior to coupling
said output datum selected to said memory for
storing results. | 15 |
| 18. A method according to claim 16 further compris-
ing the step of retrieving from a memory a sixth
and a seventh datum, and wherein in accordance
with a predetermined algorithm, said butterfly
circuit accepts as input data said sixth and sev-
enth data in lieu of said fourth and fifth data. | 20 |
| 19. A method for controlling a memory system, the
method comprising providing a structure accord-
ing to any one of claims 13-14. | 25 |

30

35

40

45

50

55

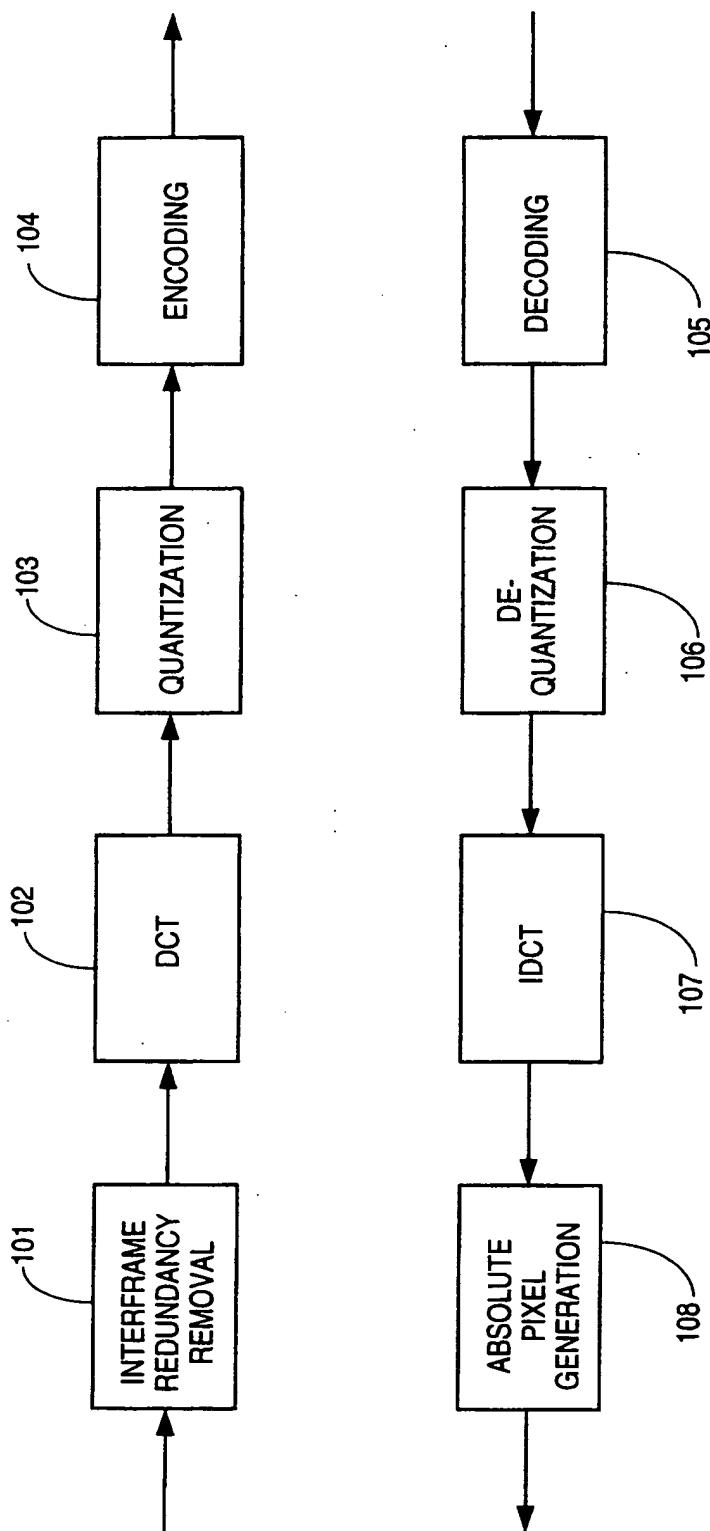


Figure 1

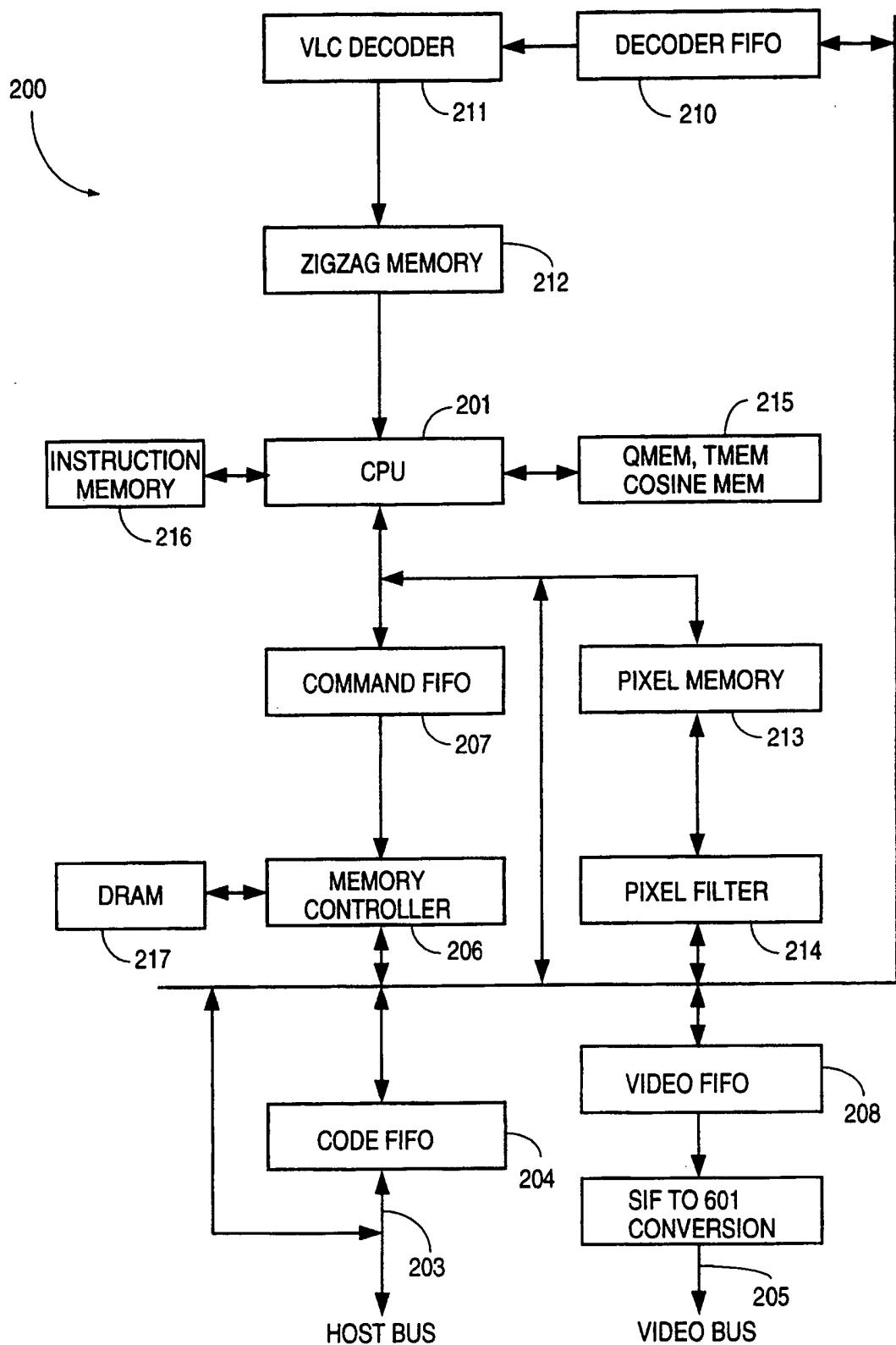
**Figure 2**

Figure 3a-1

	ZMEM	TMEM	PMEM	REG	FILE	FETCH	BUTTERFLY	MULTIPLY	REG	FILE	WRITE
	A	B	C	E	E1	E2	E3	R1	R2	R3	
1	X3	X1 ^p	-	-	-	-	-	X4	A1	B1 ^p	A1
2	X5	X7 ^p	-	B0	-	-	B _p	X0	X4	AB0	dX4
3	X1	X2 ^p	cX2	-	-	dX4	AA	X3	X0	A0 ^p	dX0
4	X6	X6 ^p	-	cX6	-	-	A _p	X5	X3	C0	dX3
5	X7	X0 ^p	iA1	-	-	-	C	dX4	X5	C1 ^p	dX5
6	X7	X4 ^p	cX0	iB1	C1	-	C _p	X1	dX4	C0 _p	C1 _p
7	X2	-	c0	c0	-	BB	BB	C1	X1	BB0	BB1
8	X3	X3 ^p	D1	BB0	AA1	D	X7	C1	X1	-	-
9	X5	X5 ^p	C0	D0	BB1	CC	dX0	AA1	X7	D0 _p	D1 _p
10	-	CC1	AA0	B1 _p	DD	BB1	dX0	AA1	CC0	CC1	dX7
11	A0 ^p	B0 _p	A1 _p	AA	A1 _p	BB1	dX0	DD0	DD0	CC1	-
12	DD1	A0 _p	C1 _p	AA _p	A1 _p	BB1	dX0	AA0	-	-	CX00
13	AAA0	D0 _p	C0 _p	BBB	C1 _p	BBB	A1 _p	B1 _p	A1 _p	AA1 _p	-
14	X6	BBB0	DC0	AA0	-	DDP	X2	C1 _p	A1 _p	-	B1 _p
15	X4	BBB1	iA1 _p	iB1 _p	dX3	DDD	X6	X2	C1 _p	DD1 _p	A1 _p
16	-	DDD0	CC0	BB0	dX5	BB _p	dX3	X6	X2	-	C1 _p
17	DDD1	D1 _p	iC1 _p	dX1	CCC	dX5	dX3	X6	BB0 _p	BB1 _p	dX2
18	-	CCC0	DD0 _p	AA0 _p	dX7	CC _p	dX1	dX5	dX3	-	dX6

19	CCC1	cX3	cX5	AA1p	DD0p	dX7	dX1	dX5	CC0p	CC1p	cX3
20	DDD0p	CC0p	BB0p	DX2	B	AA1p	DX7	DX1	-	-	-
21	DDD1p	CX1	BB1p	CCCP	A	dX2	AA1p	DX7	B0	B1	cX1
22	CCC0p	CC1p	AA0p	dX6	A	BB1p	dX2	AA1p	-	-	-
23	CCC1p	cX00	R0	B1	AAA0p	dX6	BB1p	dX2	A0	A1	-
24	AAA0p	DD1p	B0	A1	cX00	B1	dX6	BB1p	-	-	cX2
25	AAA1p	-	A0	-	BBBp	A1	B1	dX6	-	-	-
26	X0	-	-	-	-	X4	A1	B1	CX0	-	dX6
											B1

KEY TO FIGURE 3a

FIGURE 3a-1
FIGURE 3a-2

Figure 3a-2

# regs in use	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11
8	0	A0	iA1	B0	iB1	cX2	cX6	cX0	dX4	
8	1	<u>A0</u>	iA1	<u>B0</u>	iB1	cX2	cX6	cX0	dX4	
8	2		iA1		iB1	cX2	cX6	cX0	dX4	
10	3		iA1		iB1	<u>cX2</u>	<u>cX6</u>	cX0	<u>dX4</u>	
10	4		iA1		iB1			cX0		
13	5	dX3	<u>iA1</u>		<u>iB1</u>	C0	<u>C1</u>	cX0		A0p A1p
13	6	dX3	dX5			C0		<u>cX0</u>		A0p A1p
14	7	dX3	dX5	BB0	BB1	C0				A0p A1p
16	8	dX3	dX5	BB0	BB1	C0		D0	<u>D1</u>	A0p A1p
16	9	dX3	dX5	BB0	<u>BB1</u>	<u>C0</u>		<u>D0</u>		A0p A1p
16	10	dX3	dX5	BB0	dX7	CC0	<u>CC1</u>			A0p A1p
16	11	dX3	dX5	BB0	dX7	CC0		DD0	DD1	A0p <u>A1p</u>
15	12	dX3	dX5	BB0	dX7	CC0		DD0	<u>DD1</u>	<u>A0p</u>
14	13	dX3	dX5	BB0	dX7	CC0		DD0		AA0p AA1p
13	14	dX3	dX5	BB0	dX7	CC0		<u>DD0</u>		AA0p AA1p
14	15	<u>dX3</u>	dX5	BB0	dX7	CC0				AA0p AA1p
12	16		<u>dX5</u>	<u>BB0</u>	dX7	<u>CC0</u>				AA0p AA1p
12	17				dX7	dX2	dX6	BB0p	BB1p	AA0p AA1p
10	18				<u>dX7</u>	dX2	dX6	BB0p	BB1p	AA0p AA1p
11	19	<u>cX3</u>				dX2	dX6	BB0p	BB1p	AA0p <u>AA1p</u>
9	20					<u>dX2</u>	dX6	<u>BB0p</u>	BB1p	AA0p
9	21			B0	B1		dX6		<u>BB1p</u>	AA0p
8	22			B0	B1		<u>dX6</u>			<u>AA0p</u>
5	23	A0	A1	B0	<u>B1</u>					
4	24	A0	<u>A1</u>	B0		cX2				
4	25	A0		B0		cX2		cX0		
6	26	A0		B0		cX2	cX6	cX0		
6		A0		B0	iB1	cX2	cX6	cX0		
5			iA1		iB1	cX2	cX6	cX0		
6			iA1		iB1	cX2	cX6	cX0	dX4	

FIGURE 3b-1

R12	R13	R14	R15	R16	R17	R18	R19	
								0
	B0p	B1p						1
	B0p	B1p				AA0	AA1	2
	B0p	B1p			dX0	AA0	AA1	3
	B0p	B1p			dX0	AA0	AA1	4
	B0p	B1p	C0p	C1p	dX0	AA0	AA1	5
	B0p	B1p	C0p	C1p	dX0	AA0	AA1	6
	B0p	B1p	C0p	C1p	dX0	AA0	<u>AA1</u>	7
	B0p	B1p	C0p	C1p	<u>dX0</u>	AA0	dX1	8
	B0p	B1p	C0p	C1p	D0p	D1p	AA0	dX1
	B0p	<u>B1p</u>	C0p	C1p	D0p	D1p	AA0	dX1
	<u>B0p</u>		C0p	C1p	D0p	D1p	AA0	dX1
		CX00	C0p	<u>C1p</u>	D0p	D1p	AA0	dX1
		CX00	<u>C0p</u>		<u>D0p</u>	D1p	AA0	dX1
		CX00			iB1p	D1p	<u>AA0</u>	dX1
		CX00	DD0p	DD1p	<u>iB1p</u>	D1p	<u>iA1p</u>	dX1
		CX00	DD0p	DD1p	iC1p	D1p		dX1
		CX00	DD0p	DD1p	<u>iC1p</u>	<u>D1p</u>		dX1
		CX00	<u>DD0p</u>	DD1p				18
		CX00		DD1p	CC0p	CC1p		19
		CX00		DD1p	<u>CC0p</u>	CC1p		20
		CX00		DD1p		CC1p	cX1	21
		CX00		DD1p		<u>CC1p</u>		22
		<u>CX00</u>		DD1p				23
				<u>DD1p</u>				24
								25
								26

KEY TO
FIGURE 3b

FIGURE 3b-1	FIGURE 3b-2
----------------	----------------

FIGURE 3b-2

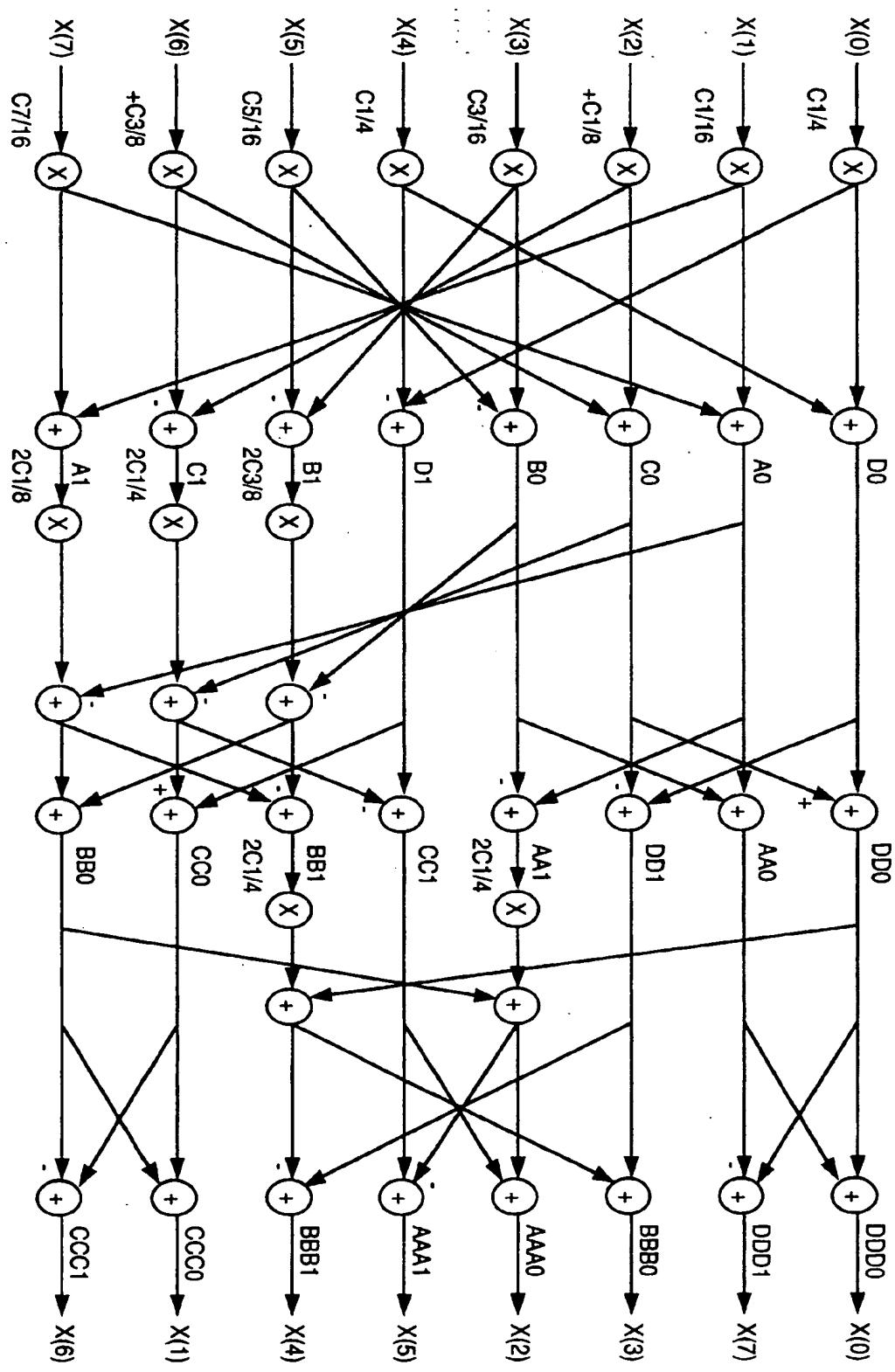


Figure 4

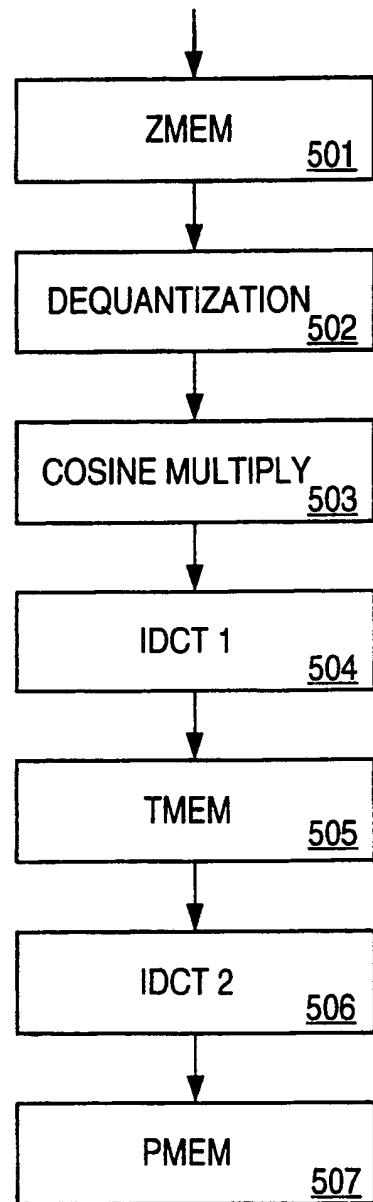


FIGURE 5

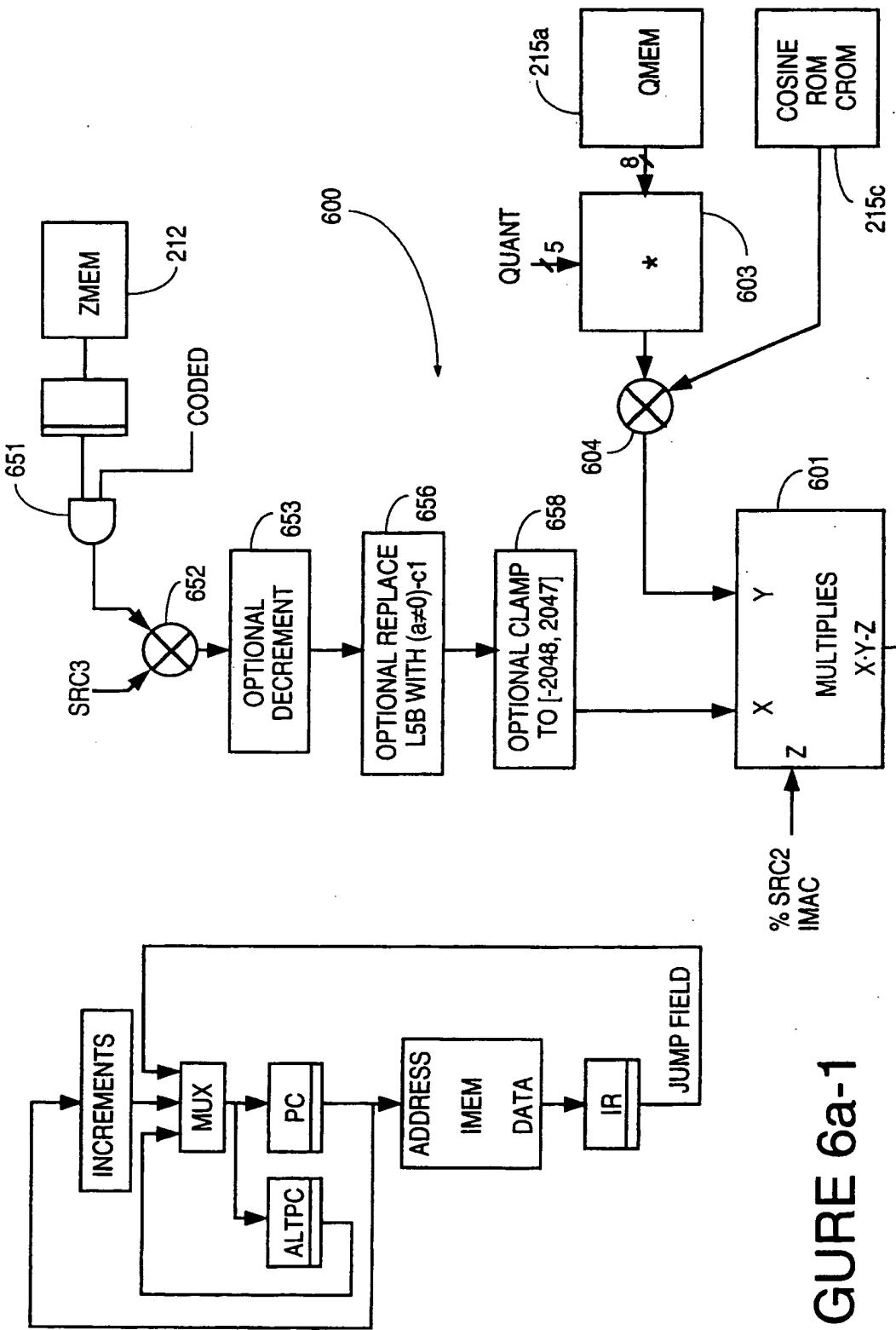
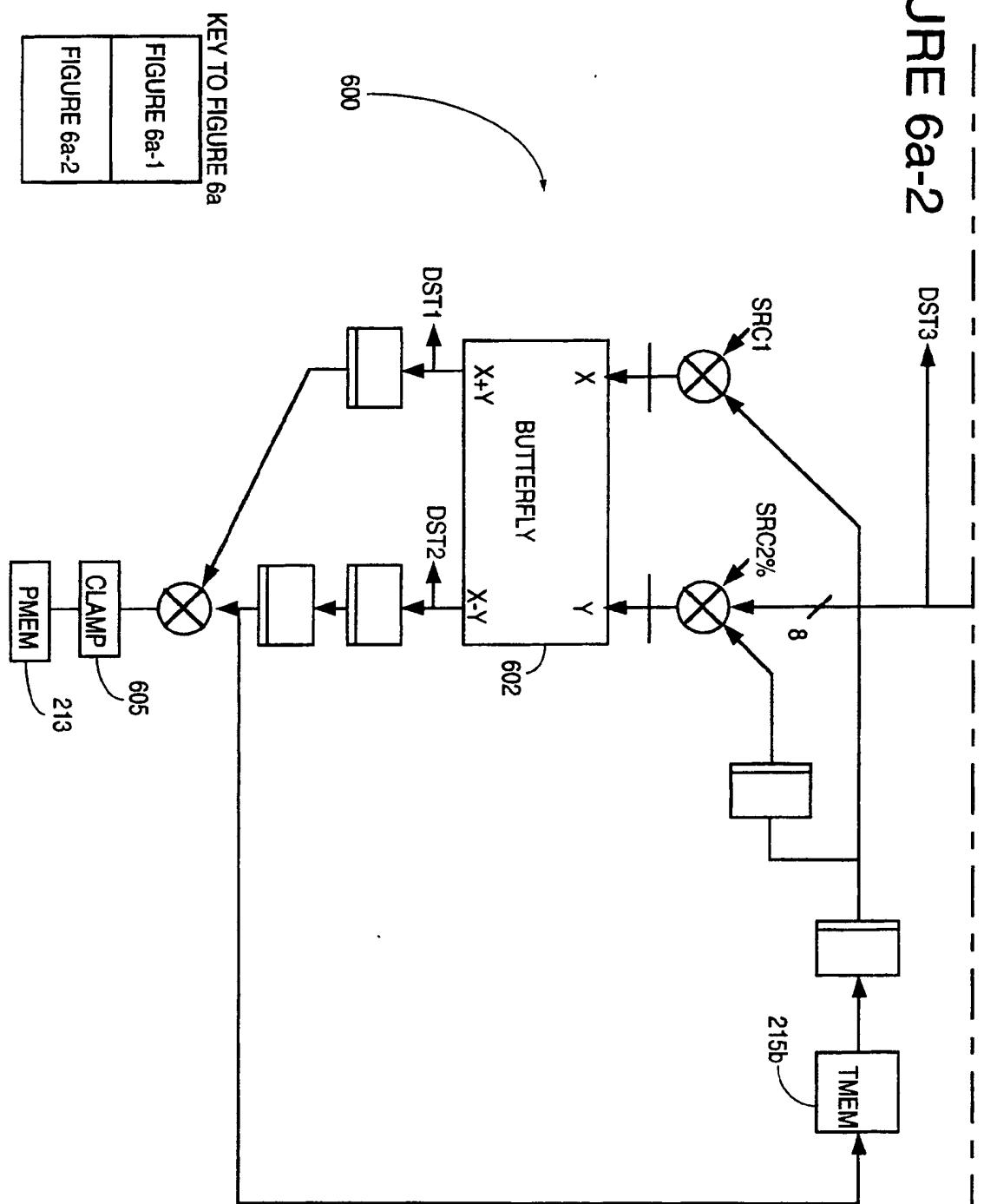


FIGURE 6a-1

FIGURE 6a-2

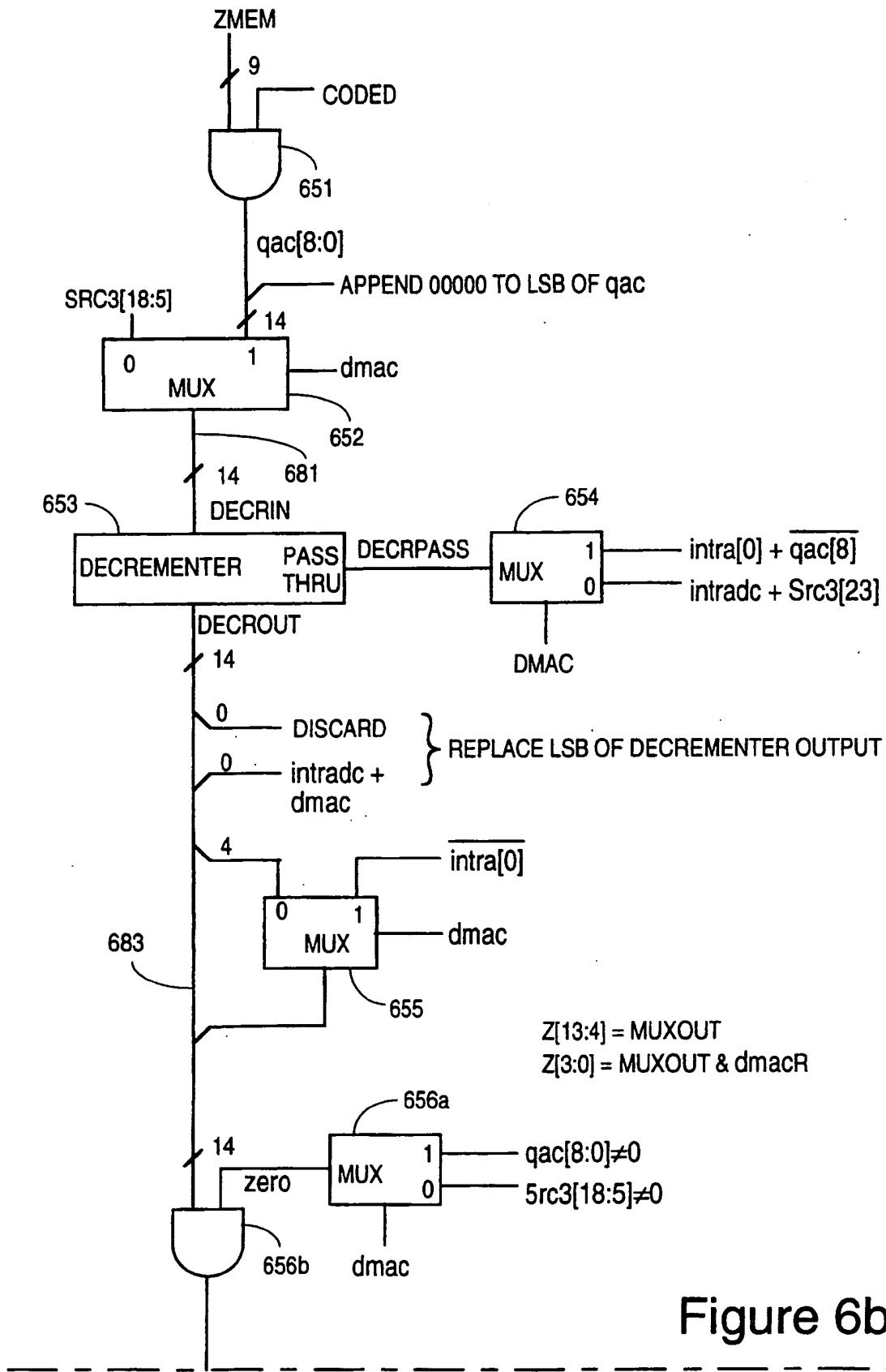
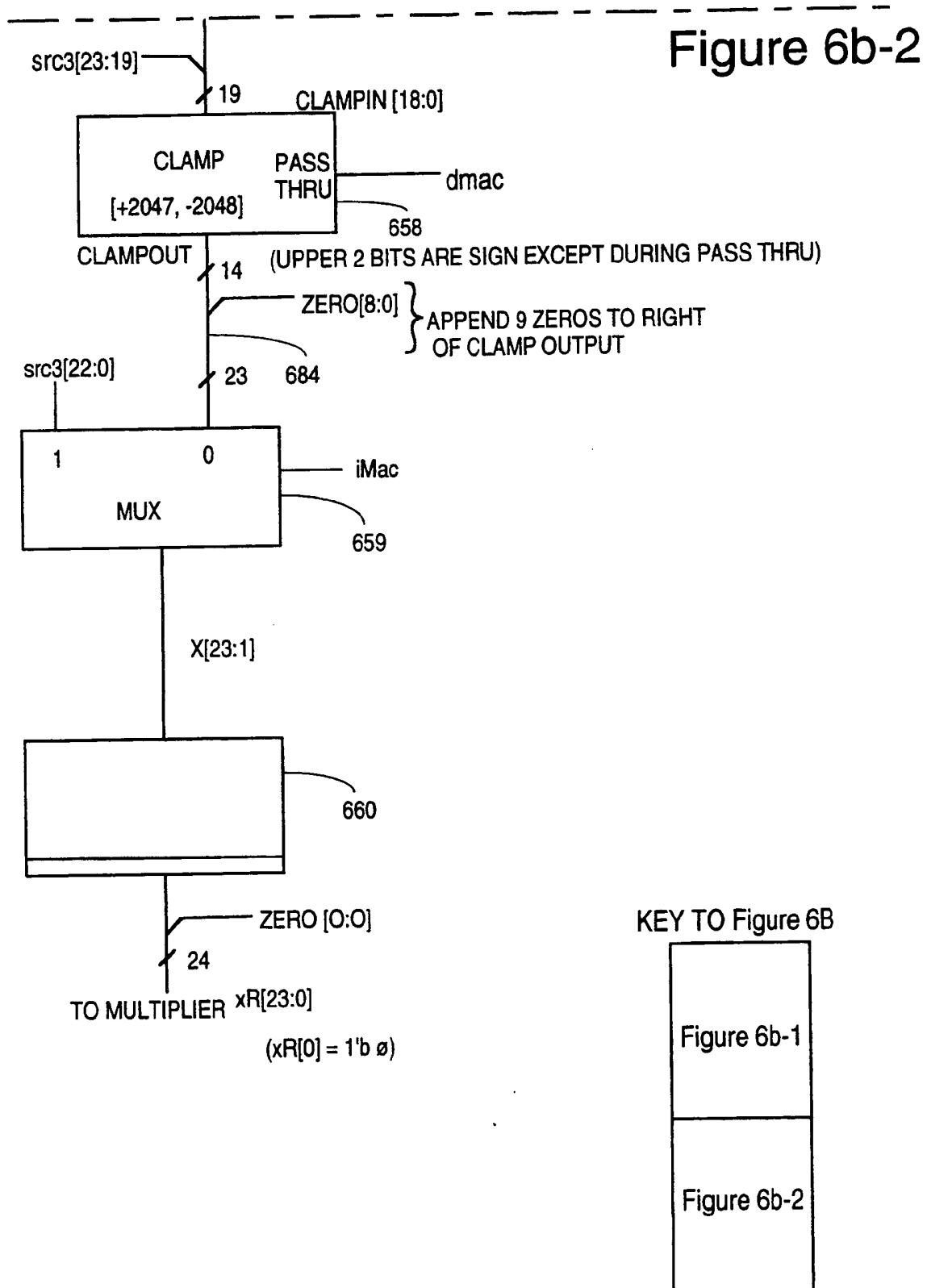


Figure 6b-1



```

uicode () {
    reg (loopcnt, r1)

    /* Registers used for dequantization and cosine multiply */
    reg(dx0,r17)          reg(dx1, r19)          reg(dx3, r2)
    reg(dx4,r9)           reg(dx5, r3)           reg(dx7, r5)
    reg(cx0,r8)           reg(cx00, r13)        reg(cx2, r6)
    reg(cx3,r2)           reg(cx6, r7)          reg(cx1, r19)

    /* Registers used for 1st pass of IDCT (ZMEM to TMEM) */
    reg(A,r2)              reg(A0, r2)           reg(A1, r3)
    reg(B,r4)              reg(B0, r4)           reg(B1, r5)
    reg(C,r6)              reg(C0, r6)           reg(C1, r7)
    reg(D,r8)              reg(D0, r8)           reg(D1, r9)
    reg(AA,r18)            reg(AA0, r18)         reg(AA1,r19)
    reg(BB,r4)             reg(BB0, r4)          reg(BB1,r5)
    reg(CC,r6)             reg(CC0, r6)          reg(CC1,r7)
    reg(DD,r8)             reg(DD0, r8)          reg(DD1,r9)

    * Registers used for 2nd pass of IDCT (TMEM to PMEM)
    reg(AP,r10)            reg(A0, r10)          reg(A1p,r11)
    reg(BP,r12)            reg(B0, r12)          reg(B1p,r13)
    reg(CP,r14)            reg(C0, r14)          reg(C1p,r15)
    reg(DP,r16)            reg(D0, r16)          reg(D1p,r17)
    reg(AAp,r10)           reg(AA0p, r10)        reg(AA1p,r11)
    reg(BBp,r8)             reg(BB0p, r8)          reg(BB1p,r9)
    reg(CCp,r16)           reg(CC0p, r16)        reg(CC1p,r17)
    reg(DDp,r14)           reg(DD0p, r14)        reg(DD1p,r15)

```

KEY TO FIGURE 7

FIGURE 7-1 	FIGURE 7-2
-----------------------	-----------------------

FIGURE 7-1

```

/*
 * DEQUANTIZATION AND IDCT OF 8x8 BLOCK
 * $ means operand bypasses register file
 * % means operand is shared between 2 ops
 * BUTTERFLYS MULTIPLIES
 */

ori (loopcnt, r0, 25
ldmac (rT, Bp, dX0, r0, r0, A0,B0
dmac (nT, AA, dX3,
dmac (rT, Ap, dX5, r0, r0, C, r4
cmac (nT, C, r0, CX2, CX6, dX4
dmac (rT, Cp, dX1, r0, r0, BB, r6
imac (nT, BB, r0, iA1, iB1, C1
dmac (BnT, D, dX7, cx0, C0
imac (rT, Dp, r0, r0, r0, AA1
cmac (BnT, CC, cx00, D1, BBO, dX0
imac (nT, DD, r0, D0, C0, BB1
imac (BwT, r0, iB1p, CC1, AA0, B1p
imac (nT, AAP, iA1p, A0p, B0p, A1p
imac (BwT, r0, iC1p, DD1, A0p, C1p
dmac (nT, DDP, dX2, D0p, C0p
dmac (wt, r0, dx6, DDO, AA0
cmac (nT, Bbp, cx3, iA1p, iB1p, dX3
cmac (wt, r0, r0, CC0, BB0, dX5
cmac (nT, CCP, CX1, D1p, iC1p, dX1
cmac (wp, r0, DD0p, AA0p, dX7
imac (BnT, B, r0, cx3, r0, AA1p
cmac (wp, r0, CX2, CC0p, BB0p, dX2
imac (BnT, A, r0, cx1, r0, BB1p
cmac (BwP, r0, CX6, CC1p, AA0p, dX6
imac (nT, cx0, iB1, cx00, r0, B1
imac (BwP, r0, iA1, DD1p, B0, A1
subi (loopcnt, loopcnt, 1
bnz (l)

/*1
/*2
/*3
/*4
/*5
/*6
/*7
/*8
/*9
/*10
/*11
/*12
/*13
/*14
/*15
/*16
/*17
/*18
/*19
/*20
/*21
/*22
/*23
/*24
/*25
/*26

Bp =b(X3p, X5p)
AA =b(A0, B0)
Ap =b(X1p, X7p)
C =b(CX2, CX6)
BB =b(X2p, X6p)
BB, r1 =b(A1, iB1)
BB, r2 =b(cx0, $cx4)
BB, r3 =b(cx2, cx6)
BB, r4 =b(cx2p, cx6p)
BB, r5 =b(A1, iB1)
BB, r6 =b(cx0, $cx4)
BB, r7 =b(cx0, $cx4)
BB, r8 =b(X0p, X4p)
BB, r9 =b(D1, $iC1)
BB, r10 =b(D0, C0)
BB, r11 =b(CC1, $IAA1)
BB, r12 =b(A0p, %BC0p)
BB, r13 =b(DD1, $IBB1)
BB, r14 =b(D0p, %C0p)
BB, r15 =b(DD0, AA0)
BB, r16 =b(iA1p, iB1p)
BB, r17 =b(CC0, BBO)
BB, r18 =b(D1p, iC1p)
BB, r19 =b(DD0p, AA0p)
BB, r20 =b(CX3, $CX5)
BB, r21 =b(CC0p, %BB0p)
BB, r22 =b(cx1, $cx7)
AAAp =b(CC1p, $iAA1p)
CX0 =b(cx00, 0)
BBBp =b(DD1p, $iBB1p)
bnz (l)

BX0 =dmac (X0)
BX3 =dmac (X3)
BX5 =dmac (X5)
$cx4 =cmac (dx4)
DX1 =dmac (X1)
$ic1 =imac (C1, C0)
DX7 =dmac (X7)
$iaa1 =imac (AA1, BB0)
CX00 =cmac (dx0)
$ibb1 =imac (BB1, AA0)
IB1p =imac (B1p, %B0p)
IA1p =imac (A1p, A0p)
IC1p =imac (C1p, %C0p)
DX2 =dmac (X2)
DX6 =dmac (X6)
CX3 =cmac (dx3)
$cx5 =cmac (dx5)
CX1 =cmac (dx1)
$cx7 =cmac (dx7)
$iaa1p =imac (AA1p, %BB0p)
CX2 =cmac (dx2)
$ibb1p =imac (BB1p, AA0p)
CX6 =cmac (dx6)
IB1 =imac (B1, B0)
IA1 =imac (A1, A0)
DX4 =dmac (x4)

*/

```

FIGURE 7-2

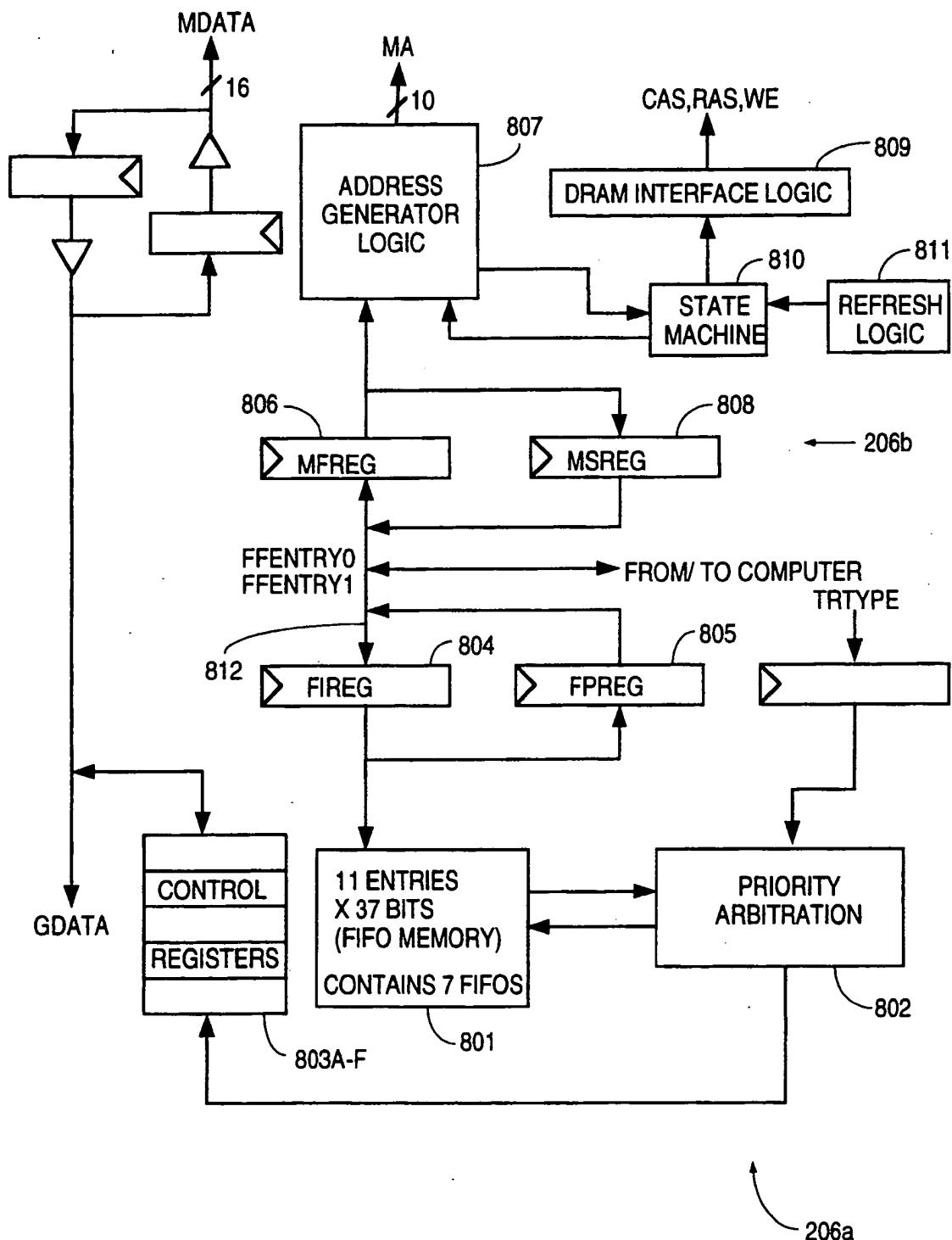


FIGURE 8

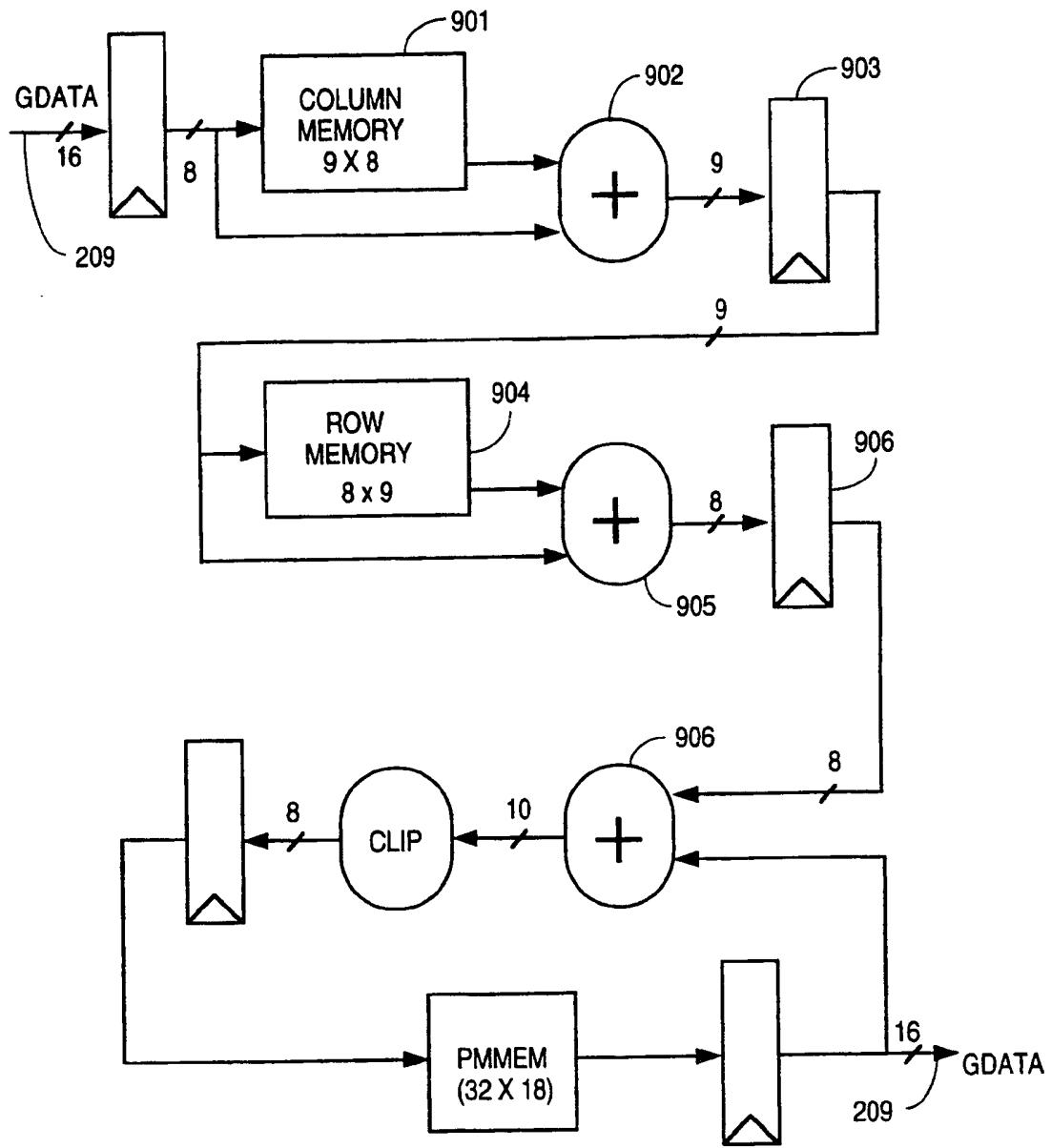


FIGURE 9

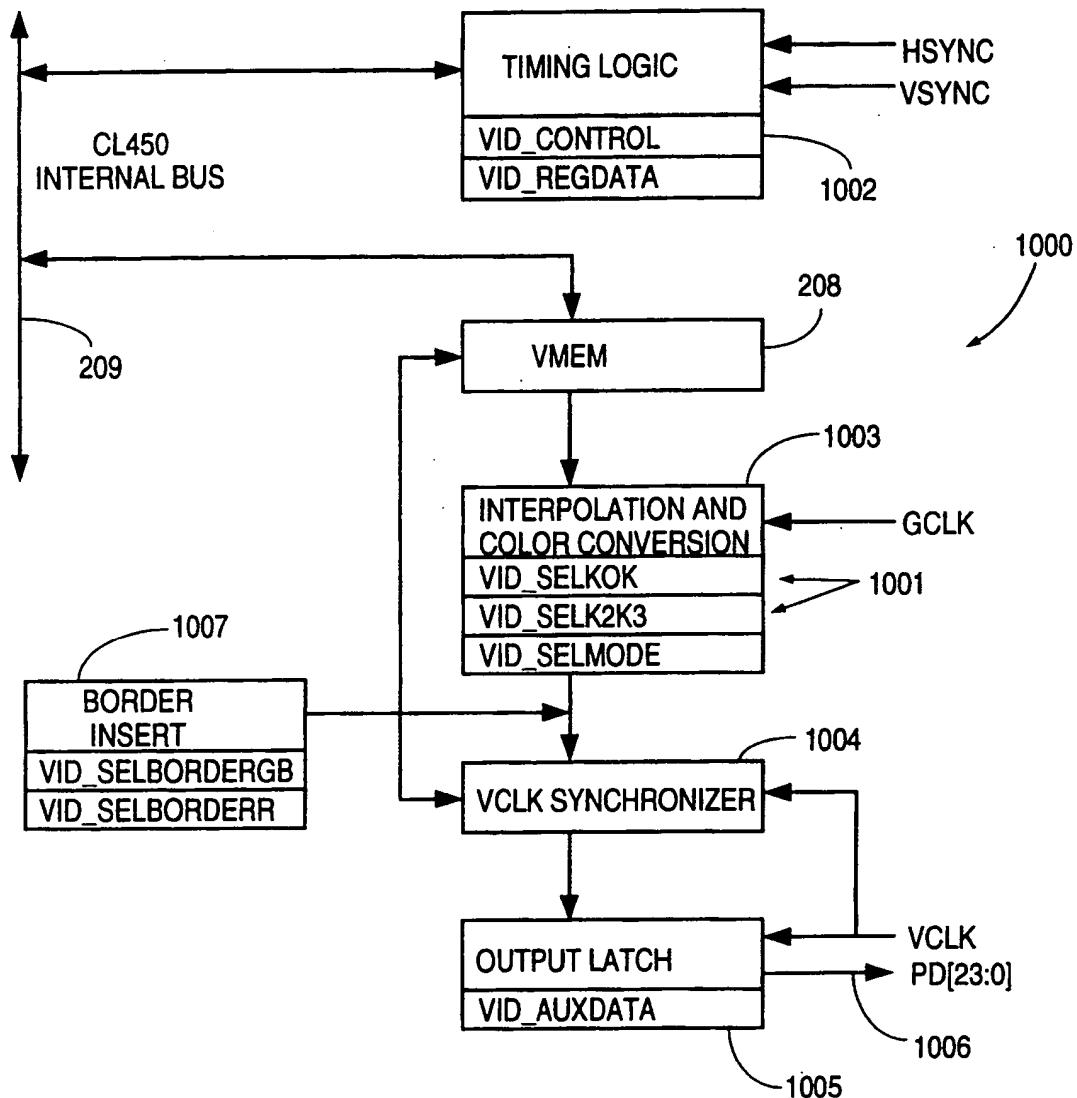
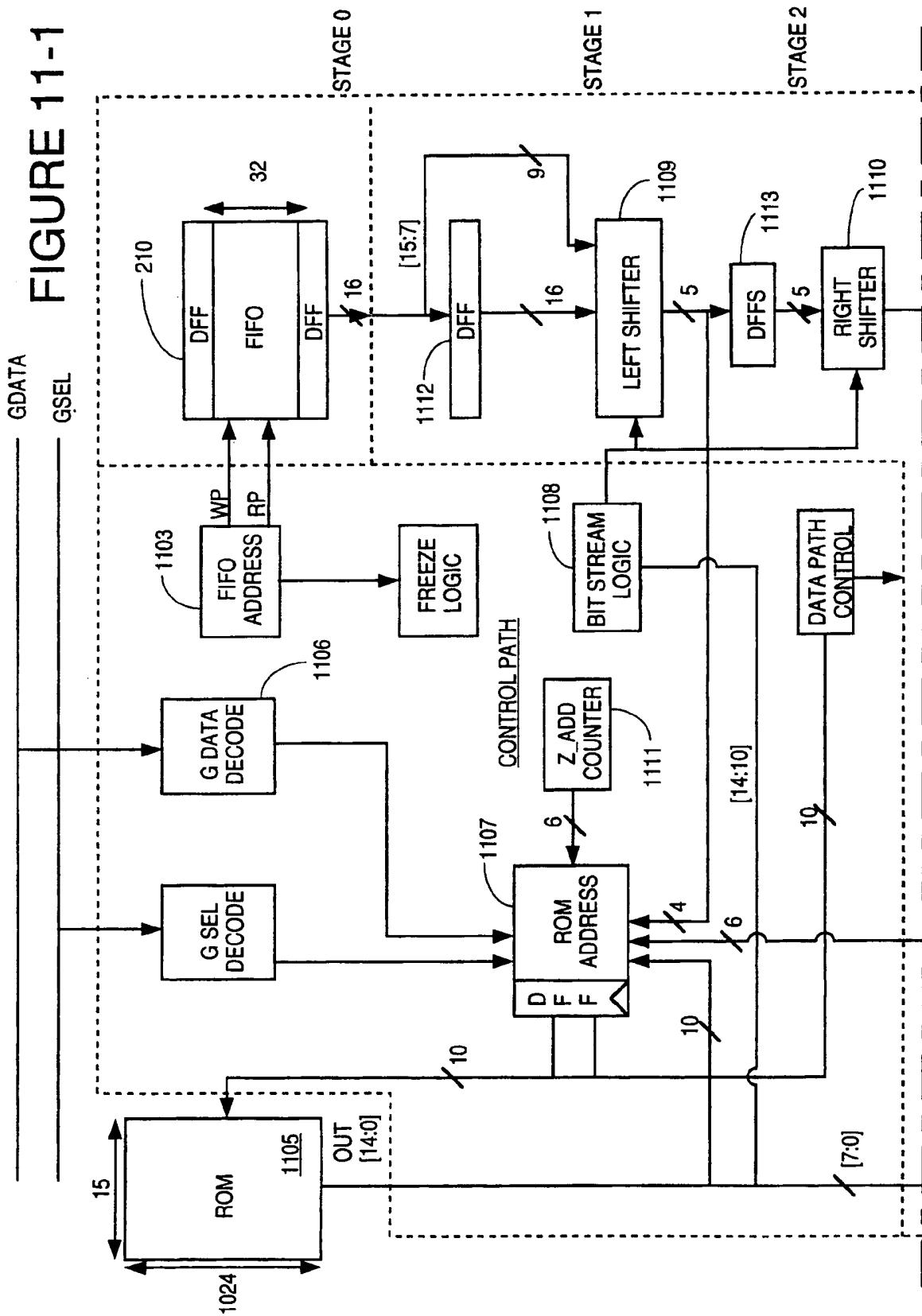


FIGURE 10

FIGURE 11-1



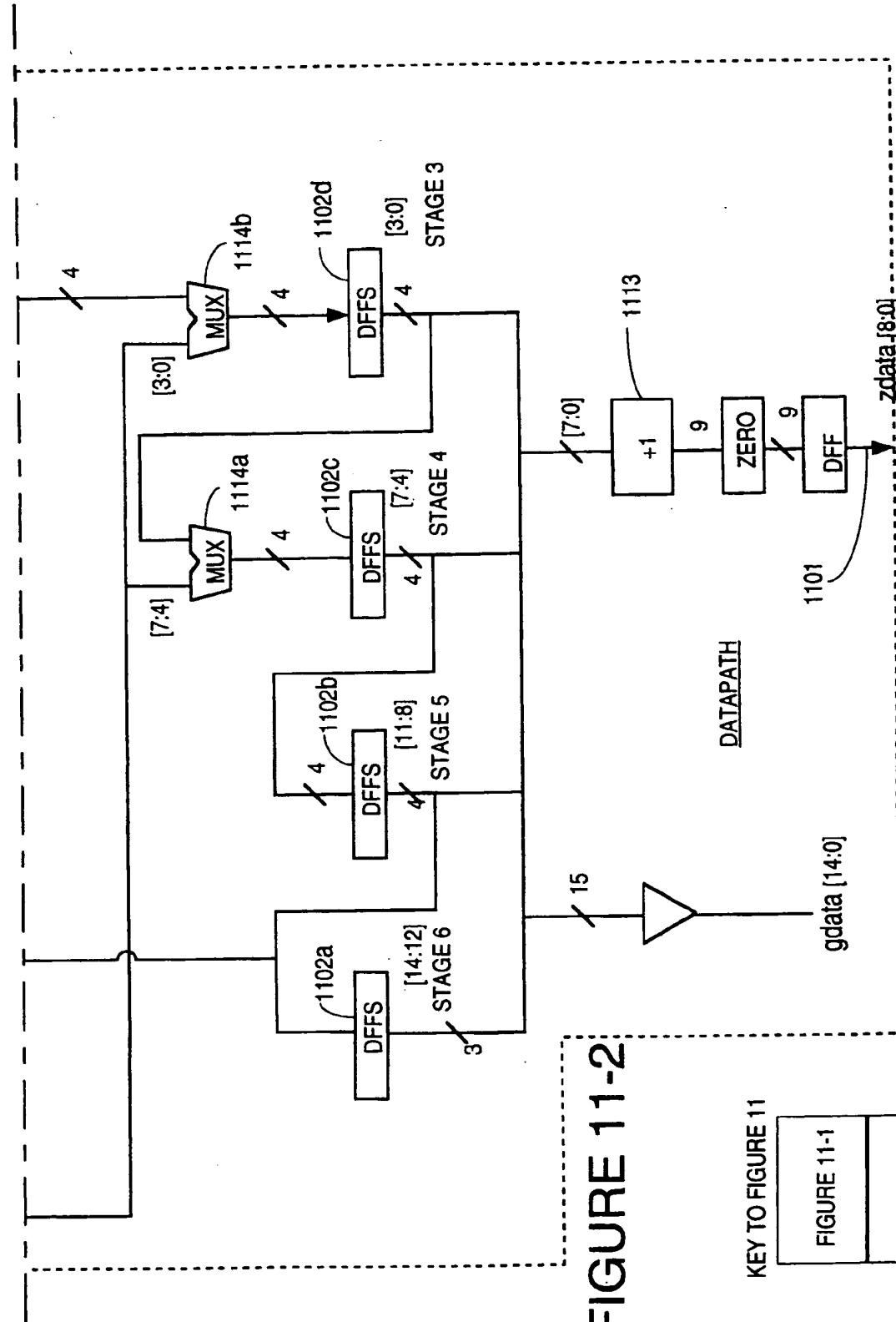


FIGURE 11-2

KEY TO FIGURE 11

FIGURE 11-1

FIGURE 11-2

THIS PAGE BLANK (USPTO)